

	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
	Brownstein M.J., Uedin T.B., Toshlyuki S., Carninci P., Prange C.,
	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
	Bosak S.A., McFawn P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
	Villalon D.K., Munzky D.M., Sodergren E.J., Lu X., Gibbs R.A.,
	Faney J.J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
	Krzywinski M.I., Skaleka U., Smailus D.E., Schmerch A., Schein J.B.,
	Jones S.J., Marra M.A.;
	"Generation and initial analysis of more than 15,000 full-length human
	and mouse cDNA sequences.";
	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
	[2]
	SEQUENCE FROM N.A.
	TISSUE=Testis;
	RA Strausberg R.;
	Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
	EMBL; BC068547; AAH68547.1; -. 736ADAD5AD444EC2 CRC64;
	SEQUENCE 688 AA; 77555 MW; 736ADAD5AD444EC2 CRC64;
	Query Match 96.6%; Score 3613; DB 2; Length 688;
	Best Local Similarity 98.5%; Pred. No. 5.2e-150;
	Matches 673; Conservative 4; Mismatches 6; Indels 0; Gaps 0;
QY	17 KREHKPKPEPQQKAPLVPPPPPPPPPPPPPPDPPTPEPEEIIIGSDSEQPADYCK 76
Db	:: ::
QY	6 KGSSSSERPEPQQAPLVPPPPPPPPPPPPPPDPPTPEPEEIIIGSDSEQPADYCK 65
QY	77 GGYHPVKIGDLFNCRGHVIRKLGWGHFSTVLWCMDQGKRFVAMKVKSQAHYTTALDE 136
Db	66 GGYHPVKIGDLFNCRGHVIRKLGNHGFSTVLWCMDQGKRFVAMKVKSQAHYTTALDE 125
QY	137 IKLLKCVRESPPDNKDMVVQLIDDPKISGMNGIHVCWFEVLGHLIKWIIKSNYQGL 196
Db	126 IKLLKCVRESDRSPDNKMVWLIDDPKI SGNNGIHVCWFEVLGHLIKWIIKSNYQGL 185
QY	197 PVRCVKSIIROVLQGLDYLHSCKI IHTDI KPENILMCVDVDAVRMAAEATEWKAGAP 256
Db	186 PVRCVKSIIROVLQGLDYLHSCKI IHTDI KPENILMCVDDAVRMAAEATEWKAGAP 245
QY	257 PPSSGAIVTAPOQKPIGKISKNNKKLKXKQKQAELEKRLQETIEELEREAEKRIIBEN 316
Db	246 PPSSGAIVTAPOQKPIGKISKNNKKLKXKQKQAELEKRLQETIEELEREAEKRIIBEN 305
QY	317 ITSAAPSNDQGEYCPEVKLTGTGLEAATAETAKONGEADEEEKEDAENIKEKDDE 376
Db	306 ITSAAPSNDQGEYCPEVKLTGTGLEAATAETAKONGEADEEEKEDAENIKEKDDE 365
QY	377 VQDELANTDPWTIESPKTNGHIENGPFSLSQOOLDEDDDEDCPNPEEYNLDPENAESDY 436
Db	366 VQDELANTDPWTIESPKTNGHIENGPFSLSQOOLDEDDDEDCPNPEEYNLDPENAESDY 425
QY	437 TYSSSYEQNGELPNGRHKIPESOPFPSTLSFGSILEPVACGSVLSEGSPLTEQEESP 496
Db	426 TYSSSYEQNGELPNGRHKIPESOPFPSTLSFGSILEPVACGSVLSEGSPLTEQEESP 485
QY	497 SHDRSRITVSASTGDL PKAKTRAADLI NVLPDPNADKI RVKTLADI CNACVWHKFTEDI 556
Db	486 SHDRSRITVSASTGDL PKAKTRAADLI NVLPDPNADKI RVKTLADI CNACVWHKFTEDI 545
QY	557 QTRQYSIEVLICAGYSTPADITWSTACMAFELATGDYLFPHSGEDYSRDEDHIAHIEL 616
Db	546 QTRQYSIEVLICAGYSTPADITWSTACMAFELATGDYLFPHSGEDYSRDEDHIAHIEL 605
QY	617 LGSI PRHFALSGKYSRFPNRNGELRHITKLPKPSLFDVLVEKYGWPHEDAQTFDLIP 676
Db	606 LGSI PRHFALSGKYSRFPNRNGELRHITKLPKPSLFDVLVEKYGWPHEDAQTFDLIP 665

QY	677	MLEWPEKRSAGACLRHPWLNS	699
DB	666	MLEWPEKRSAGACLRHPWLNS	688
RESULT 5			
P78362	ID	PRELIMINARY;	PRT; 686 AA.
AC	P78362;		
DT	01-MAY-1997	(TREMBLrel. 03, Created)	
DT	01-MAY-1997	(TREMBLrel. 03, Last sequence update)	
DT	01-WAR-2004	(TREMBLrel. 26, Last annotation update)	
DE	Serine kinase SRPK2		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
NCBI_TaxID=9606;	OX		
[1]	SEQUENCE FROM N.A.		
RN	MEDLINE=98139536; PubMed=9472028;		
RX	Wang H.Y., Lin W., Dyck J.A., Yeakley J.M., Songyang Z., Cantley L.C.,		
RA	Fu X.D.;		
RA	"SRPK2: a differentially expressed SR protein-specific kinase involved		
RT	in mediating the interaction and localization of pre-mRNA splicing		
RT	factors in mammalian cells."		
RL	J. Cell Biol. 140:737-750(1998).		
CC	-I- SIMILARITY: Belongs to the Ser/Thr protein kinase family.		
DR	EMBL; U08666; AAC05299.1; -		
DR	HSSP; Q03656; IHOW.		
DR	GO; GO:0005634; C:nucleus; TAS.		
DR	GO; GO:0004672; F:protein kinase activity; TAS.		
DR	GO; GO:0008380; P:RNA splicing; TAS.		
DR	GO; GO:0000245; P:spliceosome assembly; TAS.		
DR	InterPro; IPRO11009; Kinase like.		
DR	InterPro; IPR000719; Prot.kinase.		
DR	InterPro; IPR008271; Ser_Ch_r_kin_AS.		
DR	Pfam; PF00069; Pkinase; 1.		
DR	ProDom; PD000001; Prot.kinase; 2.		
DR	PROSITE; PS00107; PROTEIN KINASE ATP; 1.		
DR	PROSITE; PS00011; PROTEIN KINASE DOM; 1.		
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.		
DR	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.		
KW	ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.		
SQ	SEQUENCE 686 AA; 77423 MW; 1B9F16E9FEFBAB82 CRC64;		
Query Match 95.8%; Score 3586; DB 2; Length 686;			
Best Local Similarity 98.2%; Pred.No. 7.9e-149;			
Matches 671; Conservative 4; Mismatches 6; Indels 2; Gaps 2;			
QY	17	KREKHPKKPFPOOKAPLVPPPPPPPPPLDPPTPEPEEEITLGSDDEQEDPADYCK	76
DB	6	EXSSSRPFPOOKAPLVPPPPPPPPPLDPPTPEPEEEITLGSDDEQEDPADYCK	65
QY	77	GGYHPVKIGDLFNGRYHIVIRKLGGHGHSFTVWLCDWMQDKGFVAMKVKSQAHYTTETALDE	136
DB	66	GGYHPVKIGDLFNGRYHIVIRKLGGHGHSFTVWLCDWMQDKGFVAMKVKSQAHYTTETALDE	125
QY	137	IKLLKCVRSDSPDNKMDVVQLDDDFKISGMNGIHVCMPFVLGHLLKWIIKSNYQGL	196
DB	126	IKLLKCVRSDSPDNKMDVVQLDDDFKISGMNGIHVCMPFVLGHLLKWIIKSNYQGL	185
QY	197	PVRCVKSIIROVLGSLDYLHSHCKKIHTDIKPENILMCDVDAYVRRMAAEATEWQAKAP	256
DB	186	PVRCVKSIIROVLGSLDYLHSHCKKIHTDIKPENILMCDVDAYVRRMAAE-PWQKAGAP	244
QY	257	PPSGSAVSATAQQPIGKISNNKKKKUKKQKRAELLEKEKLOIEELEREAEERKIIEEN	316
DB	245	PPSGSAVSATAQQPIGKISNNKKKKUKKQKRAELLEKEKLOIEELEREAEERKIIEEN	304
QY	317	ITSAAPSNDQGEYCYPEVKLTATTGLEEAABAETAADNGEASDQEKEKAENIEKFDD	376
DB	305	ITSAAPSNDQGEYCYPEVKLTATTGLEEAABAETAADNGEASDQEKEKAENIEKFDD	364
QY	377	VDELANTDPTWIESPKNTHENGFPFSLEQOLDDDEDDEDCPNPEYNLDFNAESDY	436


```

Db      365 VQELANIDPTWIESPKTNGHIENGFFSLEQQLDDEDDDEDCPNPEYNLDEPNAESDY 424
QY      437 TVSSSYEQNGELPNGRHKIPESQPFSTSLFSGSLEPACVCSVLSEGSPLTEQESSP 496
Db      425 TVSSSYEQNGELPNGRHKIPESQPFSTSLFSGSLEPACVCSVLSEGSPLTEQESSP 484
QY      497 SHDRSRTVSASSTGDLPKAKTRAADLLVNPPLPRNADKIRVKIADLGNACWVHKHPTEDI 556
Db      485 SHDRSRTVSASSTGDLPKAKTRAADLLVNPPLPRNADKIRVKIADLGNACWVHKHPTEDI 544
QY      557 QTRQVRSIEVLGAGYSTPADISTACMAFELATGDLPEPHSGEDYSRDEDHIAHIIEL 616
Db      545 QTRQVRSIEVLGAGYSTPADISTACMAFELATGDLPEPHSGEDYSRDEDHIAHIIEL 604
QY      617 LGSIPRHPALSGKYSREFNRRGELRHITKLPWSLFDVLVSKYGMWPHEDAAQFTDFLIP 676
Db      605 LGSIPRHPALSGKYSREFNRRGELRHITKLPWSLFDVLVSKYGMWPHEDAAQFTDFLIP 664
QY      677 MLEMVPEKASAGECLRHHPWLNS 699
Db      665 MLEMVPEKASAGECLRHHPWLNS 686

```

RESULT 6

```

ID O75221 PRELIMINARY; PRT; 675 AA.
AC O75221;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-WAR-2004 (TREMBLrel. 26, Last annotation update)
DE WUGSC:H RG152G17.1b protein (Fragment).
GN Name=WUGSC:H RG152G17.1b;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rohlfing T., O'Neal D.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Waterston R.H.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Waterston R.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AC005070; RAC23141.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot kinase; 2.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
FT NON_TER 1
SQ SEQUENCE 675 AA; 76102 MW; B0FF1532AE129CEB CRC64;

```

Query Match

Best Local Similarity 95.5%; Score 3574; DB 2; Length 675;

Pred. No. 2.6e-148;

```

Matches 665; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
QY      25 PFPQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 84
Db      1 PFPQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 60
QY      85 GDLFNGRVHVIKLGWGHFSTVNLWCMDMGKRFVAMKVVKVSAQHYTETALDEIKLKVCVR 144
Db      61 GDLFNGRVHVIKLGWGHFSTVNLWCMDMGKRFVAMKVVKVSAQHYTETALDEIKLKVCVR 120
QY      145 ESDPSDPNKMVMVQLIDDFKISGMNGIHVCMVFEVLGHLLKWI IKSNTYQGLPVRVCVSI 204
Db      121 ESDPSDPNKMVMVQLIDDFKISGMNGIHVCMVFEVLGHLLKWI IKSNTYQGLPVRVCVSI 180
QY      205 IRQVLQGLDYLHCKKIIHTDIKPENILMCVDVAYVRRMAAEATEKQAGAPPPSSAVS 264
Db      181 IRQVLQGLDYLHCKKIIHTDIKPENILMCVDVAYVRRMAAEATEKQAGAPPPSSAVS 240
QY      265 TAPQOKPIGKI SKNKKKLLKKQKQAELEKRLQEI EELEREAEERKII EENITSAPSN 324
Db      241 TAPQOKPIGKI SKNKKKLLKKQKQAELEKRLQEI EELEREAEERKII EENITSAPSN 300
QY      325 DQDGEYCPEVKLKTTLGLEAAEAETAKONGEADQSEKEDAEKENI EKDEDDVDQELANI 384
Db      301 DQDGEYCPEVKLKTTLGLEAAEAETAKONGEADQSEKEDAEKENI EKDEDDVDQELANI 360
QY      385 DPTWIESPKTNGHIENGFFSLEQQLDDEDDDEDCPNPEYNLDEPNAESDYTSSSYEQ 444
Db      361 DPTWIESPKTNGHIENGFFSLEQQLDDEDDDEDCPNPEYNLDEPNAESDYTSSSYEQ 420
QY      445 FNGELPNGRHKIPESQPFSTSLFSGSLEPACVCSVLSEGSPLTEQESSPSHDSRTV 504
Db      421 FNGELPNGRHKIPESQPFSTSLFSGSLEPACVCSVLSEGSPLTEQESSPSHDSRTV 480
QY      505 SASSTGDLPKAKTRAADLLVNPPLPRNADKIRVKIADLGNACWVHKHFTEDIQTRVRSI 564
Db      481 SASSTGDLPKAKTRAADLLVNPPLPRNADKIRVKIADLGNACWVHKHFTEDIQTRVRSI 540
QY      565 EVLIGAGYSTPADISTACMAFELATGDLPEPHSGEDYSRDEDHIAHIIELGSI PRHF 624
Db      541 EVLIGAGYSTPADISTACMAFELATGDLPEPHSGEDYSRDEDHIAHIIELGKVPKY 600
QY      625 ALSGKYSREFNRRGELRHITKLPWSLFDVLVSKYGMWPHEDAAQFTDFLIPLEWVPEK 684
Db      601 AMLGKYSKEFFTRKGLRHITKLPWSLFDVLVSKYGMWPHEDAAQFTDFLIPLEWVPEK 660
QY      685 RASAGECLRHHPWLNS 699
Db      661 RASAGECLRHHPWLNS 675

```

RESULT 7

```

OS4781 PRELIMINARY; PRT; 681 AA.
ID O54781
AC O54781;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-WAR-2004 (TREMBLrel. 26, Last annotation update)
DE SRPK2.
GN Name=SRpk2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuroyanagi N., Onogi H., Wakabayashi T., Hagiwara M.;
RT "Novel SR-protein-specific kinase, SRPK2, disassembles nuclear
RT speckles."
RL Biochem. Biophys. Res. Commun. 0:0-0(1998).
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AB006036; BAA24055.1; -.
DR PIR; JC5929; JC5929.

```



```
Db 124 IKLLKCVRESDPSPDNKDMVQLIDDFKISGMNGIHVCMVFEVLGHLLKWIISNYQGL 183
Qy 197 PVRCKSIIRVQLGLDYLSKCKIIHTDIKENILMVCVDDAYVRRMAAEATEWQKAGAP 256
Db 184 PVRCKSIIRVQLGLDYLSKCKIIHTDIKENILMVCVDDAYVRRMAAEATEWQKAGAP 243
Qy 257 PPSGAVSTAPQKPIGKISKNKKLKKKKQKQAELEKRLQEIIEELREAEKRIIEEN 316
Db 244 PPSGAVSTAPQKPIGKISKNKKLKKKKQKQAELEKRLQEIIEELREAEKRIIEEN 303
Qy 317 ITSAAPSND-QDGEYCPEVKLTITGLEAAEAETAKDNGEAEDEQEKEDAENIEKORD 375
Db 304 ITSAEASGQDGEYCPEVTLKAADLEDTTEETAKDNGEVEDQEKEDAENAEKORD 363
Qy 376 DVDOELANIDPTWIESPKTNGHIENGPFSLBOOLDDEDDDEDCPNPEYNLDSEPNASD 435
Db 364 DVEOELANIDPTWIESPKANGHIENGPFSLBOOLDDEDDDEDCANPEYNLDSEPNASD 423
Qy 436 YTYSSSYEQFNGELPNGRHKIPESQPFPESTSLFSGSLPEPVACGSLVSGSPLTQEBS 495
Db 424 YTYSSSYEQFNGELPNQHKT-----SEPPTPLFSGPLPEPVACGSLVSGSPLTQEBS 478
Qy 496 PSHDRSRTVSASSTGDLPAKTRAADLLVNPDPNADKIRVKIADLGNACVWHKHFTED 555
Db 479 PSHDRSRTVSASSTGDLPAKTRAADLLVNPDPNADKIRVKIADLGNACVWHKHFTED 538
Qy 556 IOTRQVRSIEVLIGAGYSTADIWSTACMAFELATGDYLFEPHSGSDYSRDEHIAHIE 615
Db 539 IOTRQVRSIEVLIGAGYSTADIWSTACMAFELATGDYLFEPHSGSDYSRDEHIAHIE 598
Qy 616 LGSGIPRHFAKSGYSREFFNRGELRHITKLPKWSLFDVLVEKYGWPHEDAQAQFTDFLI 675
Db 599 LGSGIPRHFAKSGYSREFFNRGELRHITKLPKWSLFDVLVEKYGWPHEDAQAQFTDFLI 658
Qy 676 PMLWVPEKRSAGECLRHPLNLS 699
Db 659 PMLWVPEKRSAGECLRHPLNLS 682

RESULT 9
Q6V1X2 PRELIMINARY; PRT; 546 AA.
ID Q6V1X2
AC Q6V1X2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE SFRS protein kinase 2 isoform c.
GN Name=SRPK2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Sha J.H., Zhou Z.M., Li J.M.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AY354201; AAQ63886.1; -.
DR GO; GO:0016301; F_kinase activity; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; Pkinase; I.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYRc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
```

```
SQ SEQUENCE 546 AA; 61049 MW; 5F1DB6805A6CD79C CRC64;
Query Match 75.1%; Score 2810; DB 2; Length 546;
Best Local Similarity 98.3%; Pred. No. 4.6e-115;
Matches 525; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy 17 KREKHPKPEPOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 76
Db 6 EKSSSERPEPOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 65
Qy 77 GGYHPVKIGDLNGRHYHVRKLGWGHFSTVWLCWDMQGRFVAMKVVKSAQHYTTALDE 136
Db 66 GGYHPVKIGDLNGRHYHVRKLGWGHFSTVWLCWDMQGRFVAMKVVKSAQHYTTALDE 125
Qy 137 IKLLKCVRESDPSPDNKDMVQLIDDFKISGMNGIHVCMVFEVLGHLLKWIISNYQGL 196
Db 126 IKLLKCVRESDPSPDNKDMVQLIDDFKISGMNGIHVCMVFEVLGHLLKWIISNYQGL 185
Qy 197 PVRCKSIIRVQLGLDYLSKCKIIHTDIKENILMVCVDDAYVRRMAAEATEWQKAGAP 256
Db 186 PVRCKSIIRVQLGLDYLSKCKIIHTDIKENILMVCVDDAYVRRMAAEATEWQKAGAP 245
Qy 257 PPSGAVSTAPQKPIGKISKNKKLKKKKQKQAELEKRLQEIIEELREAEKRIIEEN 316
Db 246 PPSGAVSTAPQKPIGKISKNKKLKKKKQKQAELEKRLQEIIEELREAEKRIIEEN 305
Qy 317 ITSAAPSNDQDGEYCPEVKLTITGLEAAEAETAKDNGEAEDEQEKEDAENIEKDEDD 376
Db 306 ITSAAPSNDQDGEYCPEVKLTITGLEAAEAETAKDNGEAEDEQEKEDAENIEKDEDD 365
Qy 377 DVDOELANIDPTWIESPKTNGHIENGPFSLBOOLDDEDDDEDCPNPEYNLDSEPNASD 436
Db 366 DVDOELANIDPTWIESPKTNGHIENGPFSLBOOLDDEDDDEDCPNPEYNLDSEPNASD 425
Qy 437 YTYSSSYEQFNGELPNGRHKIPESQPFPESTSLFSGSLPEPVACGSLVSGSPLTQEBS 496
Db 426 YTYSSSYEQFNGELPNGRHKIPESQPFPESTSLFSGSLPEPVACGSLVSGSPLTQEBS 485
Qy 497 SHDRSRTVSASSTGDLPAKTRAADLLVNPDPNADKIRVKIADLGNACVWHK 550
Db 486 SHDRSRTVSASSTGDLPAKTRAADLLVNPDPNADKIRVKIADLGNACVWHK 539

RESULT 10
AAQ63886 PRELIMINARY; PRT; 546 AA.
ID AAQ63886
AC AAQ63886;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE SFRS protein kinase 2 isoform c.
GN SRPK2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primata; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Sha J.H., Zhou Z.M., Li J.M.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY354201; AAQ63886.1; -.
KW Kinase.
SQ SEQUENCE 546 AA; 61049 MW; 5F1DB6805A6CD79C CRC64;
Query Match 75.1%; Score 2810; DB 2; Length 546;
Best Local Similarity 98.3%; Pred. No. 4.6e-115;
Matches 525; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
Qy 17 KREKHPKPEPOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 76
Db 6 EKSSSERPEPOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPPP 65
```


Db 54 DDEQEDPNDYCKGYHHVKIGDLNSKYHVRKLGWGHFSTVWLAWDIQGRFVAMKV 113
QY 124 KSAHYTTALDEIKLKCVRESPPSDPNKDMVQLIDDFKISGMNGHVCWFEVLGH 183
Db 114 KSAHYTTALDEIKLRSVRNTDPPDNEMVQLDDFKISGVNGTHVCWFEVLGH 173
QY 184 LKWIISKNSYQGLPVCVKSIIROVLQGLVLSKCKIIHTDIKPENILMVCDDAYVRM 243
Db 174 LKWIISKNSYQGLPVCVKSIIROVLQGLVLSKCKIIHTDIKPENILMVCDDAYVRM 233
QY 244 AAEATEWOKAGAPPPSGSAVSTAPAKQAPKMSKNNKKKKKKQKRAELLEKRLQETEE 303
Db 234 AAEATEWOKAGAPPPSGSAVSTAPAKQAPKMSKNNKKKKKKQKRAELLEKRLQETEE 293
QY 304 LEREARKIIENITSAAPSDNDQGEYCPVVKLTGTLGEEAABEATKNGEADQEEKE 363
Db 294 MEIGPEGREEDD---DPESKPSCAP---LRQASLOEIA-----TE 330
QY 364 DAENIEKDDEDDVDQELANIDPTWIESPKTNHENGPFSLQQLDDEDD---E 416
Db 331 DIIMDTRELTSDASVEL-----NCGCLHS---SKTQPEEEDQGLQYQLLQ 376
QY 417 EDCPNPEYNLDEPNAESDYTSYSSYEQFNGELPNGRHKIPESQPFSTSLFSGSLPV 476
Db 377 EDHNN-----ANAGPETAQSMYECNG-----AESPELDQACYSNG---T 414
QY 477 ACGSVLSESGSPITEQEESPSHDSRTVSASSTGDLPAKTAADLLVNPDLPRNADKIR 536
Db 415 SGOEQLDEGELATEEHOQKTRAREQNKIKD---DKLSAGSLVNPDLPLNADKIK 471
QY 537 VKIADIGNACWVHKHFTEDIQTRVRSIEVLIGAGYSTPADITWSTACAFELATGDYLFE 596
Db 472 VKIADIGNACWVHKHFTEDIQTRVRSIEVLIGAGYSTPADITWSTACAFELATGDYLFE 531
QY 597 PHSGEDYSRDEDHIAHIIELLSGIPRHPALSGKYSREPNRGRGELRHITKLPWLSFDVL 656
Db 532 PHSGEDYSRDEDHIAHIIELLSGIPRHPALSGKYSREPNRGRGELRHITKLPWLSFDVL 591
QY 657 VEKYGWPHEDAAQFTDFLIPLEWVPEKASAGECLRHHPWL 697
Db 592 MDKYEWQBEAQFTSFDFLLPMLLEPEKRAATAADCLRHHPWL 632

RESULT 14

AAH56825 PRELIMINARY; PRT; 634 AA.
ID AAH56825; MEDLINE=22388257; PubMed=12477932;
AC AAH56825; DT 24-MAY-2004 (T-EMBLrel. 27, Created)
DT 24-MAY-2004 (T-EMBLrel. 27, Last sequence update)
DT 24-MAY-2004 (T-EMBLrel. 27, Last annotation update)
DE Hypothetical protein zgc:63789.
GN ZGC:63789.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Boak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

FAHEY J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Kryzyski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC056825; AAH56825.1; -;
KW Hypothetical protein.
SQ SEQUENCE 634 AA; 71712 MW; 95339BF0DB4B140D CRC64;

Query Match 55.1%; Score 2061; DB 2; Length 634;
Best Local Similarity 59.5%; Pred. No. 2.7e-82;
Matches 417; Conservative 71; Mismatches 135; Indels 78; Gaps 12;
QY 4 RKVLAIQARKRRPKREKHPKPEQKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPEEEILGS 63
Db 3 RKVLAIQARKRRPKREKHPKPEQKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPEEEILGS 53
QY 64 DDEQEDPNDYCKGYHHVKIGDLNSKYHVRKLGWGHFSTVWLAWDIQGRFVAMKV 123
Db 54 DDEQEDPNDYCKGYHHVKIGDLNSKYHVRKLGWGHFSTVWLAWDIQGRFVAMKV 113
QY 124 KSAHYTTALDEIKLKCVRESPPSDPNKDMVQLIDDFKISGMNGHVCWFEVLGH 183
Db 114 KSAHYTTALDEIKLRSVRNTDPPDNEMVQLDDFKISGVNGTHVCWFEVLGH 173
QY 184 LKWIISKNSYQGLPVCVKSIIROVLQGLVLSKCKIIHTDIKPENILMVCDDAYVRM 243
Db 174 LKWIISKNSYQGLPVCVKSIIROVLQGLVLSKCKIIHTDIKPENILMVCDDAYVRM 233
QY 244 AAEATEWOKAGAPPPSGSAVSTAPAKQAPKMSKNNKKKKKKQKRAELLEKRLQETEE 303
Db 234 AAEATEWOKAGAPPPSGSAVSTAPAKQAPKMSKNNKKKKKKQKRAELLEKRLQETEE 293
QY 304 LEREARKIIENITSAAPSDNDQGEYCPVVKLTGTLGEEAABEATKNGEADQEEKE 363
Db 294 MEIGPEGREEDD---DPESKPSCAP---LRQASLOEIA-----TE 330
QY 364 DAENIEKDDEDDVDQELANIDPTWIESPKTNHENGPFSLQQLDDEDD---E 416
Db 331 DIIMDTRELTSDASVEL-----NCGCLHS---SKTQPEEEDQGLQYQLLQ 376
QY 417 EDCPNPEYNLDEPNAESDYTSYSSYEQFNGELPNGRHKIPESQPFSTSLFSGSLPV 476
Db 377 EDHNN-----ANAGPETAQSMYECNG-----AESPELDQACYSNG---T 414
QY 477 ACGSVLSESGSPITEQEESPSHDSRTVSASSTGDLPAKTAADLLVNPDLPRNADKIR 536
Db 415 SGOEQLDEGELATEEHOQKTRAREQNKIKD---DKLSAGSLVNPDLPLNADKIK 471
QY 537 VKIADIGNACWVHKHFTEDIQTRVRSIEVLIGAGYSTPADITWSTACAFELATGDYLFE 596
Db 472 VKIADIGNACWVHKHFTEDIQTRVRSIEVLIGAGYSTPADITWSTACAFELATGDYLFE 531
QY 597 PHSGEDYSRDEDHIAHIIELLSGIPRHPALSGKYSREPNRGRGELRHITKLPWLSFDVL 656
Db 532 PHSGEDYSRDEDHIAHIIELLSGIPRHPALSGKYSREPNRGRGELRHITKLPWLSFDVL 591
QY 657 VEKYGWPHEDAAQFTDFLIPLEWVPEKASAGECLRHHPWL 697
Db 592 MDKYEWQBEAQFTSFDFLLPMLLEPEKRAATAADCLRHHPWL 632

RESULT 15
Q81Y12

This Page Blank (uspto)


```
Db 424 TYSSYEONFELPNGQHTSEETP-----LFGLEFPVACGSVISGSPLTEQESSP 478
Qy 497 SHDRSRTVSAAGTGLPKAKTRAADLLVNPDPNADKIRVKIADLGNACVWHKHFTEDI 556
Db 479 SHDRSRTVSAAGTGLPKTKTFAADLLVNPDPNADKIRVKIADLGNACVWHKHFTEDI 538
Qy 557 QTRQVRSIEVLIGAGYSTPADIWSTACMAFELATGDYLFEPHSGSDYSRDEBHTAHIEL 616
Db 539 QTRQVRSIEVLIGAGYSTPADIWSTACMAFELATGDYLFEPHSGSDYSRDEBHTAHIEL 598
Qy 617 LGSIPRHFAISGKYSREFNRRGELRHITKLPWSLFDVLVEKYGWPHEDAAQFTDFLIP 676
Db 599 LGSIPRHFAISGKYSREFNRRGELRHITKLPWSLFDVLVEKYGWPHEDAAQFTDFLIP 658
Qy 677 MLEVPPEKRASAGECLRHPLWNS 699
Db 659 MLEVPPEKRASAGECLRHPLWNS 681

RESULT 2
S45337
serine protein kinase SRPK1 - human
C:Species: Homo sapiens (man)
C:Date: 10-Dec-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S45337
R:Gui, J.F.; Lane, W.S.; Fu, X.D.
Nature 369, 678-682, 1994
A:Title: A serine kinase regulates intracellular localization of splicing factors in the
A:Reference number: S45337; PMID:94268559; PMID:8208298
A:Accession: S45337
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-655 <GUI>
A:Cross-references: UNIPROT:Q12890; GB:U09564; NID:G507212; PIDN:AAA20530.1; PID:G507213

Query Match 55.3%; Score 2070.5; DB 2; Length 655;
Best Local Similarity 58.8%; Pred. No. 1.1e-87;
Matches 419; Conservative 84; Mismatches 133; Indels 77; Gaps 14;
Qy 4 RKVLAIAQARKR--PKREKHPKPPQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPEEIL 61
Db 3 RKVLAIAQARKR--PKREKHPKPPQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPEEIL 49
Qy 62 GSDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCMDQGRFVAMK 121
Db 50 GSDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCMDQGRFVAMK 109
Qy 122 VVKSQAQHTETALDIBIKLLKCVRESDPDPNKMVQVLLDDPKISGMNGIHVCMVFEVLG 181
Db 110 VVKSQAQHTETALDIBIRLLKSVRNSDPNDPNEMVQVLLDDPKISGVNGTHICMVFEVLG 169
Qy 182 HLLKWIITKSNYQGLPVRCKVSIIRQVQLGLDYLHSCKIIHTDIPKPNILMCDVDDAYVR 241
Db 170 HLLKWIITKSNYQGLPVCVKIIIOVQLGLDYLHSCRIIHTDIPKPNILSVNEQYIR 229
Qy 242 RMAAEATEWQKAGAPPPSGSAVSTAPQPKIGIKSNKKKKLKKQKQKQAELEKRLQEI 301
Db 230 RMAAEATEWQKAGAPPPSGSAVSTAPQPKPADKMSKNKKLKKQKQKQAELEKRMQEI 289
Qy 302 ELEBEAEARKIIEENITSAAPNDQGEYCPVVKLTGTEEAABAEATAKNGEABDQEE 361
Db 290 EEMEKES-----GPGQKRP-----KQEESESVDRPLTENPN 328
Qy 362 KEDAKENIEKDDVDQELANIDPTWIESPKTNGHIE--NGPFSLEQQDDDD-- 416
Db 329 KLE-ESSTIGQQTLMERDTEGGAELN-----CNGVIE--VINTQNSNNTLRHK 377
Qy 417 EDCPNPEEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFESTSLFSGSLRPV 476
Db 378 EDLHNANDCDVNLQAESSFL-----SLPNGDSS--TSQETDSDCTPITSEVSDTM 425
Qy 477 ACGSVLSEGSPLTEQEES-----PSHDSRSTVSASSTGDLF-KAKTRAADLLVNP 526
```

```
Db 426 VCQSSSTVGGSFSQHIISQIQESTRAEIPCEDEQ---EOEHNGFLDNKKGSTAGNFLVNP 482
Qy 527 LDPNADKIRVKIADLGNACVWHKHFTEDIQTRQVRSIEVLIGAGYSTPADIWSTACMAF 586
Db 483 LEPKNAEKLVKIADLGNACVWHKHFTEDIQTRQVRSIEVLIGAGYSTPADIWSTACMAF 542
Qy 587 ELATGDYLFEPHSGSDYSRDEBHTAHIELLGSIPRHFAISGKYSREFNRRGELRHITK 646
Db 543 ELATGDYLFEPHSGSEETRDEBHTAHIELLGVKPRKLIIVAGKYSKEFTFKGDLKHITK 602
Qy 647 LKPNSLFDVLVEKYGWPHEDAAQFTDFLIPMLEVMPPEKRASAGECLRHPLWNS 699
Db 603 LKPNGLFEVLVEKYEWSQEEAAGFTDFLIPMLELIPKRRATAABCLRHPLWNS 655

RESULT 3
JC5930
serine/arginine-rich protein-specific kinase (EC 2.-.-.-) 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004
C:Accession: JC5930
R:Kuroyanagi, N.; Onogi, H.; Wakabayashi, T.; Hagiwara, M.
Biochem. Biophys. Res. Commun. 242, 357-364, 1998
A:Title: Novel SR-protein-specific kinase, SRPK2, disassembles nuclear speckles.
A:Reference number: JC5929; MUID:98113357; PMID:9446799
A:Accession: JC5930
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-648 <KUR>
A:Cross-references: UNIPROT:O70193; UNIPROT:O70551; UNIPROT:Q99J73
A:Experimental source: brain
C:Comment: This enzyme regulates the disassembly of the serine/arginine-rich proteins in
C:Keywords: transferase

Query Match 55.0%; Score 2058; DB 2; Length 648;
Best Local Similarity 57.7%; Pred. No. 3.9e-87;
Matches 414; Conservative 88; Mismatches 123; Indels 92; Gaps 13;
Qy 4 RKVLAIAQARKR--PKREKHPKPPQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPEEIL 61
Db 3 RKVLAIAQARKR--PKREKHPKPPQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPEEIL 49
Qy 62 GSDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCMDQGRFVAMK 121
Db 50 GSDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCMDQGRFVAMK 109
Qy 122 VVKSQAQHTETALDIBIKLLKCVRESDPDPNKMVQVLLDDPKISGMNGIHVCMVFEVLG 181
Db 110 VVKSQAQHTETALDIBIRLLKSVRNSDPNDPNEMVQVLLDDPKISGVNGTHICMVFEVLG 169
Qy 182 HLLKWIITKSNYQGLPVRCKVSIIRQVQLGLDYLHSCKIIHTDIPKPNILMCDVDDAYVR 241
Db 170 HLLKWIITKSNYQGLPVCVKIIIOVQLGLDYLHSCRIIHTDIPKPNILSVNEQYIR 229
Qy 242 RMAAEATEWQKAGAPPPSGSAVSTAPQPKIGIKSNKKKKLKKQKQKQAELEKRLQEI 301
Db 230 RMAAEATEWQKAGAPPPSGSAVSTAPQPKPADKMSKNKKLKKQKQKQAELEKRMQEI 289
Qy 302 ELEBEAEARKIIEENITSAAPNDQGEYCPVVKLTGTEEAABAEATAKNGEABDQEE 361
Db 290 EEMEKES-----GPGQKRP-----KQEESESVDRPLTENPN 323
Qy 362 KEDAKENIEKDDVDQELANIDPTWIESPKTNGHIE--NGPFSLEQQDDDD-- 415
Db 324 KMTQEK---LEBSNIGQ---DQTLTERGEGGAPEINCNGVIGVNVYNSNNETLR 375
Qy 416 EDCPNPEEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFESTSLFSGSL 474
Db 376 HKEDLHNANDCDVHTLKQEPSFLNS-----NG-----DSSPSQDTSCT 415
Qy 475 PVACGSVLSEGS-----PLTQEES-----PSHDSRSTVSASSTGDLF-KAKTRAADL 522
```

Db 416 PTASETMVCSAEQSLTRQDITQLEESIRADTPSPGDEQPNGLDSDS-----KGKFSAGNF 471

Qy 523 LVNPLDPRNADKIRVKIADLGNACVWVKHFTEDIQTRQVRSIEVLIGAGYSTPADIWSTA 582

Db 472 LINPLEPKNAELQVKIADLGNACVWVKHFTEDIQTRQVRSIEVLIGAGYSTPADIWSTA 531

Qy 583 CMAFELATGDYLFEPHSGSDYRDEDHIAHIELLGSIPRHFALSGKYSREFNRRGELR 642

Db 532 CMAFELATGDYLFEPHSGSDYRDEDHIAHIELLGSIPRHFALSGKYSREFNRRGELR 591

Qy 643 HITKLPWLSFDVLVEKYGMPHEDAAQFTDFLIPMLEMVPKRSAGSECLRHPLWS 699

Db 592 HITKLPWLGLEVLVEKYGMPEAAAGFTDFLLPMLEMPEKRAATAECLRHPLWS 648

RESULT 4

S28282

hypothetical protein B0464.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 12-Mar-1993 #sequence_revision 12-Mar-1993 #text_change 09-Jul-2004

C:Accession: S28282

R:Kershaw, J.

submitted to the EMBL Data Library, December 1992

A:Reference number: S28278

A:Accession: S28282

A:Molecule type: DNA

A:Residues: 1-1087 <KER>

A:Cross-references: UNIPROT:Q03563; EMBL:Z19152

C:Genetics:

A:Introns:

C:Keywords: serine/threonine-specific protein kinase

Query Match 36.7%; Score 1374; DB 2; Length 1087;

Best Local Similarity 46.7%; Pred. No. 1.3e-55;

Matches 305; Conservative 68; Mismatches 158; Indels 122; Gaps 13;

Qy 53 PPEPEEILGSDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCDWM 112

Db 481 PMPDGEQ-LGSDDEQEDPRDYKRGYHPVNIQGVFNARYHVIRKLGWGHFSTVWLAWDT 539

Qy 113 QGKRFVAMKVKSAAHYTETALDEIKLLKCVRESPPSPDKMVKQVQLIDDFKISGMNGIH 172

Db 540 QDKRFVAMKIVKSAEHTAALDEIKLLSVRSADPNDIGCHVKVQLLDEFTVTGNGQH 599

Qy 173 VCMVEVLGHLLKWIISNYOGLPVRVKSIIROVLQGLDYLSKCKIIHTDIPKENTL 232

Db 600 VAMVEVLGCNLLKIIRSNYRGLHLEQVRKICRQVLEALGYMHEKCGIIHTDIPENVL 659

Qy 233 MCVDDAYVRRMAAEATEWQKAGAPPSPGSASVTAPO---QKPIGKISKNKKKKKKKK 289

Db 660 ITMSREEIKMAQHAVVARKNM-KMSGSAVSTAPDHLVYKMAQENNTKNNKKKKKKK 718

Qy 290 QALLEKRLQELIEELEREARKIIEENITSAAPSDQDGEYCPEVKLTGTEEAEEAET 349

Db 719 QREKLEAELAGLEGLKMDANG---LQAYNAPLITNIGKYSMCMNNGTLEENP----- 771

Qy 350 AKDNGEADQERKEDAEKENIEKDEDDVDQELANIDPTWIESPKTNGHTEGPFPSLEQ 409

Db 772 ---NASQVEDVTMEDVTNEN-----GNRNKVEIRSPDRFDTLTTPFS----- 811

Qy 410 DDEDDDEEDCPNPEY--NLDEPNABSDTYSSSYEQFNGELPNGRHKIPESQPFESTS 467

Db 812 -----DPESKFGDLASPSAE---YLSS-----PMSQLP----- 836

Qy 468 LFGSLLEPVACGVLSEGSPLTEQESSPSHDSRTVSASTGDLPAKTRAADLLVNL 527

Db 837 ---PGILPAP-----PV 846

Qy 528 DPNRAD---KIRVKIADLGNACVWVKHFTEDIQTRQVRSIEVLIGAGYSTPADIWSTA 584

Db 847 GPNIGDPYCIDVKIADLGNACVWVNHHTDDIQTRQVRALEVLIGSGYGPADIWSTAC 906

Qy 585 AFELATGDYLFEPHSGSDYRDEDHIAHIELLGSIPRHFALSGKYSREFNRRGELRH 644

Db 907 AFELATGDYLFEPHOGDNYSRDEDLAHISELLGAIPPSIYKKGRWRFFPHKNGHLLHI 966

Qy 645 TKLKPWLSFDVLVEKYGMPHEDAAQFTDFLIPMLEMVPKRSAGSECLRHPL 697

Db 967 HQLKPWSLVLEVRQYKWSHEDAQQFESFLRPMDFDQEKRTAKIALKHPFL 1019

RESULT 5

F88556

protein B0464.5a [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C:Accession: F88556

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolo

A:Reference number: A75000; MUID:9069613; PMID:9851916

A>Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_el

A>Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; an

A:Accession: F88556

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1093 <STO>

A:Cross-references: UNIPROT:Q03563; GB:chr_III; PIDN:CAA79540.1; PID:g3873796; GSPDB:GN

C:Genetics:

A:Gene: B0464.5a

A:Map position: 3

Query Match 36.7%; Score 1374; DB 2; Length 1093;

Best Local Similarity 46.4%; Pred. No. 1.4e-55;

Matches 304; Conservative 64; Mismatches 143; Indels 144; Gaps 13;

Qy 53 PPEPEEILGSDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCDWM 112

Db 481 PMPDGEQ-LGSDDEQEDPRDYKRGYHPVNIQGVFNARYHVIRKLGWGHFSTVWLAWDT 539

Qy 113 QGKRFVAMKVKSAAHYTETALDEIKLLKCVRESPPSPDKMVKQVQLIDDFKISGMNGIH 172

Db 540 QDKRFVAMKIVKSAEHTAALDEIKLLSVRSADPNDIGCHVKVQLLDEFTVTGNGQH 599

Qy 173 VCMVEVLGHLLKWIISNYOGLPVRVKSIIROVLQGLDYLSKCKIIHTDIPKENTL 232

Db 600 VAMVEVLGCNLLKIIRSNYRGLHLEQVRKICRQVLEALGYMHEKCGIIHTDIPENVL 659

Qy 233 MCVDDAYVRRMAAEATEWQKAGAPPSPGSASVTAPO---QKPIGKISKNKKKKKKKK 289

Db 660 ITMSREEIKMAQHAVVARKNM-KMSGSAVSTAPDHLVYKMAQENNTKNNKKKKKKK 718

Qy 290 QALLEKRLQELIEELEREARKIIEENITSAAPSDQDGEYCPEVKLTGTEEAEEAET 349

Db 719 QREKLEAELAGLEGLKMDANG---LQAYNAPLITNIGKYSMCMNNGTLEENP----- 748

Qy 350 AKDNGEADQER--KEDAEKENIEKDEDDVDQELANIDPTWIESPKTNGHTEGPFPSLEQ 407

Db 749 ELENFNASQVEDVTMEDVTNEN-----GNRNKVEIRSPDRFDTLTTPFS----- 793

Qy 408 QLDDDDDEEDCPNPEY--NLDEPNABSDTYSSSYEQFNGELPNGRHKIPESQPFESTS 465

Db 794 -----DPESKFGDLASPSAE---YLSS-----PMSQLP----- 818

Qy 466 TSLFGSLLEPVACGVLSEGSPLTEQESSPSHDSRTVSASTGDLPAKTRAADLLVNL 525

Db 819 ----PGILPAP----- 826

Qy 526 PLDPNAD---KIRVKIADLGNACVWVKHFTEDIQTRQVRSIEVLIGAGYSTPADIWSTA 582

Db 827 PVPNGIDPYCIDVKIADLGNACVWVNHHTDDIQTRQVRALEVLIGSGYGPADIWSTA 866

Qy 583 CMAFELATGDYLFEPHSGSDYRDEDHIAHIELLGSIPRHFALSGKYSREFNRRGELR 642

Db 887 CMAFELATGDYLFEPHOGDNYSRDEDLAHISELLGAIPPSIYKKGRWRFFPHKNGHLL 946

QY 516 K----- 516
Db 597 KYPNSNDVYKFEKDINKPPIYCDMENHLIHPREALRLHLYMKNKNDISNNTMDLGN 556
QY 517 -----TRAADLLVNLPLDR--NADKIRVKIADLGNACVHVKHFTEDIQTRYRSI 564
Db 657 QNSHKVYINTDEGEYCIRPDSPVYVYHEKSCVKICDLGNLWIDBSRYAEIQTRYRAP 716
QY 565 EVLIGAGYSTPADINSTACWAFELATGDYLFEPHSGEDYSRDEDHIAHIELLGSIPRHF 624
Db 717 EVILKSGFNETADISFACWVPELVTDGLFNPQKGRYDKNEEHLSPFIEVLGNIPKGM 776
QY 625 ALSGKYSRFFNRG--ELRHITKLKPSLFDVLVEKYGWPHEDAAQFTDFLIPMLEMVE 683
Db 777 IDAGYNHSHYFNKNVYLNKIRNIKKYGLYKILKYKYNLPKEKISPLCSFLIPMLSVDPQ 836
QY 684 KRASAGECLRHHPWLN 698
Db 837 TRPSAVTMLQHPWLN 851

RESULT 12
C84553
probable protein kinase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: C84553
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umavam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84553
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-440 <STO>
A:Cross-references: UNIPROT:Q9SHL5; GB:AE002093; NID:G4914374; PIDN:AAD32910.1; GSPDB:GN
C:Genetics:
A:Gene: At2g17530
A:Map position: 2
C:Superfamily: human protein kinase clkl; protein kinase homology

Query Match 22.4%; Score 838; DB 2; Length 440;
Best Local Similarity 30.3%; Pred. No. 1.5e-31;
Matches 194; Conservative 78; Mismatches 133; Indels 236; Gaps 11;

QY 63 SDDEQEDPAD-YCKGGYHPVKIGDLF-NGRYHVIRKLGWGHFSTVWLCMDMQGRFVAM 120
Db 7 SGSEDDDEGFDAYRKGGYHVRIGDQFAGGRVIAQRKLGWGFSTVWLAIDRTSNYVAL 66

QY 121 KVKSAQHYTETALDEIKLLKCVRESDDPNKDMVVQLIDDFKISGMNGIHVCMVFEVL 180
Db 67 KIQSALQPAQAALHEILLQAAADGPE--NTKCVIRLIDDFKHAGPNGQHLGVLEFL 124

QY 181 GHLLKWIKSNYQGLPVRCSIIQVLOGLDYLSKCKIIHTDIKPENILMC--VDDA 238
Db 125 GDSLLRLIKNYKGMELSKVREICCKILTGDLHRELGMHSDLKPENILLCSTID-- 182

QY 239 YVRRMAAEATEKQAGAPPSGSAVSTAPQOKPIGKISKNK-KKXKKKQKQAELEKR 297
Db 183 -----PAKOPIRSLTP-----ILEKPEGNQNGTSTWNLEKLLKRAKAAAKISGR 230

QY 298 LOEIELEAREARKIIEENITSAAPSDODGEYCPVEVLTGTEBAAAEATAKNGEAE 357
Db 231 -----RVSIVLS----- 238

QY 358 DOEKEKDAKENIEKEDDDVQOELANIDPTWIESPKTNGHIENGPFSLQQLDDEDEE 417
Db 239 -----ETPKN-----KRNLDDGID-- 252

QY 418 DCPNPEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFPEFSTLSGSLPVA 477

Db 253 ----- 252
QY 478 CGSVLSEGSPLTEQESSPSHRSRTVSASSTGDLPLKAKTRAADLLVNLPLDRNADKIRV 537
Db 253 -----MRC 255
QY 538 KIADLGNACVHVKHFTEDIQTRYRSIEVLIGAGYSTPADINSTACWAFELATGDYLFEP 597
Db 256 KVVDFGNGCWADNKFAEIQTRYRAPEVILSGYSYSDVMMSFACTAFELATGDMFLPAP 315
QY 598 HSGEDYSRDEDHIAHIELLGSIPRHFALSGYKSYRFFNRGELRHITKLKPSLFDVLV 657
Db 316 KEGNGYGEDDHIALMWELLGKMPKIAIGGARSXDYFDRHGDGLKRIRLKYWPLDRLLI 375
QY 658 EKYGPHFHDAQAQFTDFLIPMLEMVEPKASASAGECLRHHPWLN 698
Db 376 DKYKLPEAEAREFADFCLPIMDFAPEKRPPTAQOCLQHPWLN 416

RESULT 13
T04655
protein kinase homolog F8D20.10 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T04655
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes
submitted to the Protein Sequence Database, July 1998
A:Reference number: Z15381
A:Accession: T04655
A:Molecule type: DNA
A:Residues: 1-438 <BEV>
A:Cross-references: UNIPROT:O81783; EMBL:AL031135
A:Experimental source: cultivar Columbia; BAC clone F8D20
C:Genetics:
A:Map position: 4
A:Introns: 61/3; 326/3; 342/3
A:Note: F8D20.10
C:Superfamily: human protein kinase clkl; protein kinase homology

Query Match 22.3%; Score 834.5; DB 2; Length 438;
Best Local Similarity 31.3%; Pred. No. 2.2e-31;
Matches 202; Conservative 71; Mismatches 129; Indels 243; Gaps 13;

QY 62 GSDDEQEDPADYCKGGYHPVKIGDLF-NGRYHVIRKLGWGHFSTVWLCMDMQGRFVAM 120
Db 8 GSEGEEGFDA-YRKGGYHVRIGDPSGGRVIAQRKLGWGFSTVWLAIDTLTSTYVAL 66

QY 121 KVKSAQHYTETALDEIKLLKCVRESDDPNKDMVVQLIDDFKISGMNGIHVCMVFEVL 180
Db 67 KIQSALQPAQAALHEIFLSAAADGD-LDKTK-CVRLIDHFHSGPNGQHLGVLEFL 124

QY 181 GHLLKWIKSNYQGLPVRCSIIQVLOGLDYLSKCKIIHTDIKPENILMC--VDDA 238
Db 125 GDSLLRLIKNYQGLKLVKREICRCILTGDLHRELGMHSDLKPENILLCSTID-- 182

QY 239 YVRRMAAEATEKQAGAPP-----PSGSAVSTAPQOKPIGKISKNKKKKKQKQO-AEL 293
Db 183 -----PAKOPVRSGLTPLEKPEGNANGGASTWNLI-----EKKLKRRAKRAVAKI 228

QY 294 LEKRLQETEEELEREARKIIEENITSAAPSDODGEYCPVEVLTGTEBAAAEATAKON 353
Db 229 SERRVSMV-----TG--EASSTK-- 245

QY 354 GEADQEEKDAKENIEKEDDDVQOELANIDPTWIESPKTNGHIENGPFSLQQLDDEDE 413
Db 246 -----EKSIDGID----- 253

QY 414 DDEDCPNPEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFPEFSTLSFSGSL 473
Db 254 ----- 253

QY 474 EPVACGSVLSEGSPLTEQESSPSHRSRTVSASSTGDLPLKAKTRAADLLVNLPLDRNAD 533


```

QY 283 LKKKQKRAELLEKLQLEIELEERAERKIIIEENITSAAPSDQDQGYCPEVKLKTGLE 342
Db 205 -----VSSENV----- 210
QY 343 EAAEAETAKNGEAEQOEKEDAEKENIEKEDDDVDQELANIDPTWIESPKTNCHIENG 402
Db 211 ----- 210
QY 403 FSLEQOQDDEDEDCPNPEYNILDEPNABSDYTYSSSYEOFNGELPNGRHKIPESOPP 462
Db 211 -----KLPON--- 215
QY 463 EFSTSLFSGSLFPVACGVLSEGSPLTEQESSPSHDSRTVVSASSTGDLPAKTRAADL 522
Db 216 -----KSAANETHFRC-----LPKSSA----- 233
QY 523 LVNPLDPRNADKIRVKIADLGNACVWHKHFTEDIOIOTQYRSIEVLIGAGYSTPADIWSTA 582
Db 234 -----IKLIDFGSTVCNRIHHSIVQTRHYRSPFVILGLWSYQCDLWSIG 279
QY 583 CMAFELATGYLFEPHSGEDYSRDEHIAHIIELGSIPIPHFALSGKYSREFFNRRG--- 639
Db 280 CILFELCTGEALFQTHD-----NLEHLAMMERALGPLPEHMTKASRGAEKYFRGCGRL 333
QY 640 -----ELRHITKLPWLSFDVLVEKYCWPHED--AAQFTDPLIPMLEWVPEK 684
Db 334 NWPEGANSRESIRAVKRLDRLKD-----MVSU-----HVDNTRSRPADLILGLLAYDPSE 383
QY 685 RASAGECLRHPPWLS 699
Db 384 RLTAENEALDHPFEKS 398

```

Search completed: January 19, 2005, 15:41:35
Job time : 49 secs

...uge Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3742	100.0	699	9	US-09-759-359A-2	Sequence 2, Appli
2	3742	100.0	699	14	US-10-207-973-2	Sequence 2, Appli
3	3742	100.0	699	16	US-10-699-676-2	Sequence 2, Appli
4	3626	96.9	688	17	US-10-618-941-102	Sequence 102, Appli
5	3614	96.6	715	15	US-10-425-114-54382	Sequence 54382, 2
6	3586	95.8	686	10	US-08-981-397A-22	Sequence 23, Appli
7	2070.5	55.3	655	14	US-10-116-275-142	Sequence 142, Appli
8	1882	50.3	533	14	US-10-353-690-48	Sequence 48, Appli
9	1882	50.3	533	17	US-10-723-860-1549	Sequence 1549, Appli
10	1374	36.7	698	14	US-10-369-493-5839	Sequence 5839, Appli
11	1364.5	36.5	782	14	US-10-369-493-5838	Sequence 5838, Appli
12	1076.5	28.8	550	15	US-10-425-114-49483	Sequence 49483, Appli
13	1071.5	28.6	547	15	US-10-424-599-215731	Sequence 215731, Appli

```
QY 181 GHLLKWIISKYQGLPVRCKVSIIRQVLQGLDYLSKCKIIHTDIKPNILMCVDDAYV 240
Db 181 GHLLKWIISKYQGLPVRCKVSIIRQVLQGLDYLSKCKIIHTDIKPNILMCVDDAYV 240
QY 241 RMAAEATEWOKAGAPPPSGSAVSTAPQOKPIGKISKNKKKLKKQKRAELLEKRLQ 300
Db 241 RMAAEATEWOKAGAPPPSGSAVSTAPQOKPIGKISKNKKKLKKQKRAELLEKRLQ 300
QY 301 IEELEREAEKRIIEENITSAAPSNDDQGEYCPEVKLKTTLGLEAAAEATAKNDGEAE 360
Db 301 IEELEREAEKRIIEENITSAAPSNDDQGEYCPEVKLKTTLGLEAAAEATAKNDGEAE 360
QY 361 EKEDAENIEKDEDDVDQELANIDPTWIESPKTNHNGIENGPFSLQQLDDDDDEEDCP 420
Db 361 EKEDAENIEKDEDDVDQELANIDPTWIESPKTNHNGIENGPFSLQQLDDDDDEEDCP 420
QY 421 NPEEYNLDEPNAESDYTYSSSYEQNGELPNGRHKIPESQPFSTSLFSGSLEPVACGS 480
Db 421 NPEEYNLDEPNAESDYTYSSSYEQNGELPNGRHKIPESQPFSTSLFSGSLEPVACGS 480
QY 481 VLSEGSPLTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNPLDPRNADKIRVKIA 540
Db 481 VLSEGSPLTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNPLDPRNADKIRVKIA 540
QY 541 DLGNACWVHKHFTEDIQTRQYRSIEVLIGAGYSTPADIWSTACMAFELATGDLPEPHSG 600
Db 541 DLGNACWVHKHFTEDIQTRQYRSIEVLIGAGYSTPADIWSTACMAFELATGDLPEPHSG 600
QY 601 EDYSRDEDHIAHIELLGSIPRHFAISGKYSREFFNRGELRHITKLPWSLFDVLVVKY 660
Db 601 EDYSRDEDHIAHIELLGSIPRHFAISGKYSREFFNRGELRHITKLPWSLFDVLVVKY 660
QY 661 GWPHEHDAQAQTFDLIPMLEMVEPKASAGECLRHHPWLS 699
Db 661 GWPHEHDAQAQTFDLIPMLEMVEPKASAGECLRHHPWLS 699
```

RESULT 2

```
US-10-207-973-2
; Sequence 2, Application US/10207973
; Publication No. US20030175927A1
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/10/207,973
; CURRENT FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Human
US-10-207-973-2
```

```
Query Match 100.0%; Score 3742; DB 14; Length 699;
Best Local Similarity 100.0%; Pred. No. 9,6e-198;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSSRKVLTAQARRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPEE 60
Db 1 MSSRKVLTAQARRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPEE 60
QY 61 LGSDDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCWDMQGRFVAM 120
Db 61 LGSDDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCWDMQGRFVAM 120
QY 121 KVVSAQHYTETALDEIKLLKCVRESDDPSDPNKMVVQLIDDFKISGMNGIHCVMFVYL 180
Db 121 KVVSAQHYTETALDEIKLLKCVRESDDPSDPNKMVVQLIDDFKISGMNGIHCVMFVYL 180
```

```
QY 181 GHLLKWIISKYQGLPVRCKVSIIRQVLQGLDYLSKCKIIHTDIKPNILMCVDDAYV 240
Db 181 GHLLKWIISKYQGLPVRCKVSIIRQVLQGLDYLSKCKIIHTDIKPNILMCVDDAYV 240
QY 241 RMAAEATEWOKAGAPPPSGSAVSTAPQOKPIGKISKNKKKLKKQKRAELLEKRLQ 300
Db 241 RMAAEATEWOKAGAPPPSGSAVSTAPQOKPIGKISKNKKKLKKQKRAELLEKRLQ 300
QY 301 IEELEREAEKRIIEENITSAAPSNDDQGEYCPEVKLKTTLGLEAAAEATAKNDGEAE 360
Db 301 IEELEREAEKRIIEENITSAAPSNDDQGEYCPEVKLKTTLGLEAAAEATAKNDGEAE 360
QY 361 EKEDAENIEKDEDDVDQELANIDPTWIESPKTNHNGIENGPFSLQQLDDDDDEEDCP 420
Db 361 EKEDAENIEKDEDDVDQELANIDPTWIESPKTNHNGIENGPFSLQQLDDDDDEEDCP 420
QY 421 NPEEYNLDEPNAESDYTYSSSYEQNGELPNGRHKIPESQPFSTSLFSGSLEPVACGS 480
Db 421 NPEEYNLDEPNAESDYTYSSSYEQNGELPNGRHKIPESQPFSTSLFSGSLEPVACGS 480
QY 481 VLSEGSPLTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNPLDPRNADKIRVKIA 540
Db 481 VLSEGSPLTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNPLDPRNADKIRVKIA 540
QY 541 DLGNACWVHKHFTEDIQTRQYRSIEVLIGAGYSTPADIWSTACMAFELATGDLPEPHSG 600
Db 541 DLGNACWVHKHFTEDIQTRQYRSIEVLIGAGYSTPADIWSTACMAFELATGDLPEPHSG 600
QY 601 EDYSRDEDHIAHIELLGSIPRHFAISGKYSREFFNRGELRHITKLPWSLFDVLVVKY 660
Db 601 EDYSRDEDHIAHIELLGSIPRHFAISGKYSREFFNRGELRHITKLPWSLFDVLVVKY 660
QY 661 GWPHEHDAQAQTFDLIPMLEMVEPKASAGECLRHHPWLS 699
Db 661 GWPHEHDAQAQTFDLIPMLEMVEPKASAGECLRHHPWLS 699
```

RESULT 3

```
US-10-799-676-2
; Sequence 2, Application US/10799676
; Publication No. US20040157297A1
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043DIV II
; CURRENT APPLICATION NUMBER: US/10/799,676
; CURRENT FILING DATE: 2004-03-15
; PRIOR APPLICATION NUMBER: 10/207,973
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/759,359
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-799-676-2
```

```
Query Match 100.0%; Score 3742; DB 16; Length 699;
Best Local Similarity 100.0%; Pred. No. 9,6e-198;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSSRKVLTAQARRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPEE 60
Db 1 MSSRKVLTAQARRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPEE 60
QY 61 LGSDDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCWDMQGRFVAM 120
Db 61 LGSDDDEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCWDMQGRFVAM 120
```

Qy	121	KVKSQAQYHTETALDEIKLLKCVRESDSPDNKMVQVLIIDDFKISGNGIHVCMVFEVL	180
Db	121	KVKSQAQYHTETALDEIKLLKCVRESDSPDNKMVQVLIIDDFKISGNGIHVCMVFEVL	180
Qy	181	GHLLKWIINKNYOGLPVRCKVSIIRVQLGLDYLHSCCKIIHTDIKPENILCMVDDAVV	240
Db	181	GHLLKWIINKNYOGLPVRCKVSIIRVQLGLDYLHSCCKIIHTDIKPENILCMVDDAVV	240
Qy	241	RMMAAEATEWOKAGAPPSGSAVSTAPOKPIGKISKNKKKLLKKQKRQAELEKRLQE	300
Db	241	RMMAAEATEWOKAGAPPSGSAVSTAPOKPIGKISKNKKKLLKKQKRQAELEKRLQE	300
Qy	301	IEELEREAEKRIIBENITSAAPSNDDGGEYCPEVXLKTTGLLEAAEAETAKONGAEADQE	360
Db	301	IEELEREAEKRIIBENITSAAPSNDDGGEYCPEVXLKTTGLLEAAEAETAKONGAEADQE	360
Qy	361	EKEDAEKENIEKEDDDVQDELANDPTWIESPKTNGHIENGPFSLQEQLDDDEDDDEDCP	420
Db	361	EKEDAEKENIEKEDDDVQDELANDPTWIESPKTNGHIENGPFSLQEQLDDDEDDDEDCP	420
Qy	421	NPEEYNLDPEPNAESDYTTSSSYEQFNGELPNGRHKIPESQPEFSTLSFGSLEPVA CGS	480
Db	421	NPEEYNLDPEPNAESDYTTSSSYEQFNGELPNGRHKIPESQPEFSTLSFGSLEPVA CGS	480
Qy	481	VLSGSPLTEQEBSPPSHDRSRTVSASTGDLPKAKTRAADILVNPLDPNADKIRVKIA	540
Db	481	VLSGSPLTEQEBSPPSHDRSRTVSASTGDLPKAKTRAADILVNPLDPNADKIRVKIA	540
Qy	541	DLGNACVWHKFTEDIOPTQVRSIEVLIGAGYSTPADIWSTACMAFELATGDYLEPEHSG	600
Db	541	DLGNACVWHKFTEDIOPTQVRSIEVLIGAGYSTPADIWSTACMAFELATGDYLEPEHSG	600
Qy	601	EDYSRDEDHIAHIIELLGSIIPHFALSGKISREFFNRRGELRHITKLPWSLFDVLVVKY	660
Db	601	EDYSRDEDHIAHIIELLGSIIPHFALSGKISREFFNRRGELRHITKLPWSLFDVLVVKY	660
Qy	661	GWPHEDAAQFTDFILPMLVEMPKEKASAGECLRHHPWLNS	699
Db	661	GWPHEDAAQFTDFILPMLVEMPKEKASAGECLRHHPWLNS	699

```

RESULT 4
US-10-618-941-102
; Sequence 102, Application US/10618941
; Publication No. US20040197792A1
; GENERAL INFORMATION:
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: CAENEPEL, SEAN
; TITLE OF INVENTION: NOVEL KINASES
; FILE REFERENCE: 034536-0321
; CURRENT APPLICATION NUMBER: US/10/618,941
; CURRENT FILING DATE: 2003-07-15
; PRIOR APPLICATION NUMBER: 60/395,632
; PRIOR FILING DATE: 2002-07-15
; NUMBER OF SEQ ID NOS: 143
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 102
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-618-941-102

```

Query Match 96.9%; Score 3626; DB 17; Length 688;
Best Local Similarity 98.8%; Pred. NO. 2.3e-191;
Matches 675; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy	17	KREKHPKKPEPQQAFLVPPPPPPPPPPPLDPTTPEPEEILGSDDEQEDPADYCK	76
	:	:	
Db	6	EKSSSERPEPQQAFLVPPPPPPPPPPPLDPTTPEPEEILGSDDEQEDPADYCK	65
Qy	77	GGYHPVKIGDLFNGRYHYIVIRKLGWGHFSTWVLCDWQMGKRFAMKVKSAQHYTTALDE	136

Db	66	GGYHPVKIGDLNFNGRYHVIRKLGWGHSTVWLCWDMQGKRFVAMKVKSQAQHYTETALDE	125
Qy	137	IKLLKCVRESDPSPDNKDMVQVLIDDPKFISGMNGIGHVCMVFEVLGHLLKWIISKNSYQGL	196
Db	126	IKLLKCVRESDPSPDNKDMVQVLIDDPKFISGMNGIGHVCMVFEVLGHLLKWIISKNSYQGL	185
Qy	197	PVRCVKSIIRQVLQGLDYLHSHKCKIIHTDIIKPNILMCVDDAYVRRMAAEATEWOKAGAP	256
Db	186	PVRCVKSIIRQVLQGLDYLHSHKCKIIHTDIIKPNILMCVDDAYVRRMAAEATEWOKAGAP	245
Qy	257	PPSGSVSTPAQQPIGKISKNKKKLLKKKQKQAELEKRLQETEELEREAEERKIIEN	316
Db	246	PPSGSVSTPAQQPIGKISKNKKKLLKKKQKQAELEKRLQETEELEREAEERKIIEN	305
Qy	317	ITSAAPSNDDQGEYCPEVKLKTTLGLEAAAEATAKNGEAEDEQEEKEDAENKIEKDEDD	376
Db	306	ITSAAPSNDDQGEYCPEVKLKTTLGLEAAAEATAKNGEAEDEQEEKEDAENKIEKDEDD	365
Qy	377	VDQELANIDPTWIIESPKNTHIENGPSLEOQLDDEDDDEDCPNPEYNLDPNAESDY	436
Db	366	VDQELANIDPTWIIESPKNTHIENGPSLEOQLDDEDDDEDCPNPEYNLDPNAESDY	425
Qy	437	TYSSSYEQFNGELPNGRHKIPESOFPPEPSTSLFGSLSPVACGSLSEGSPLTEQEBESSP	496
Db	426	TYSSSYEQFNGELPNGRHKIPESOFPPEPSTSLFGSLSPVACGSLSEGSPLTEQEBESSP	485
Qy	497	SHDRSRTVSASSTGDLPKAKTRAADLLVNPDPNADKIRVKIADLGNACWVHKHFTEDI	556
Db	486	SHDRSRTVSASSTGDLPKAKTRAADLLVNPDPNADKIRVKIADLGNACWVHKHFTEDI	545
Qy	557	QTRQYRSIEVLIGAGYSTPADINWSTACWAPELATGDYLFEPHSGEDYGRDEDHIAHIEL	616
Db	546	QTRQYRSIEVLIGAGYSTPADINWSTACWAPELATGDYLFEPHSGEDYGRDEDHIAHIEL	605
Qy	617	LGS1PRHFA1SGKYSREFFNRGRBELRH1TKLKPWSLFDVLVEKYGWPHEDAAQFTDFLIP	676
Db	606	LGS1PRHFA1SGKYSREFFNRGRBELRH1TKLKPWSLFDVLVEKYGWPHEDAAQFTDFLIP	665
Qy	677	MLEMVPKRSASAGECLRHPLWLS	699
Db	666	MLEMVPKRSASAGECLRHPLWLS	688

RESULT 5

```

US-10-425-114-54382
; Sequence 54382, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 54382
; LENGTH: 715
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3063-134-E5_FLI.pep
US-10-425-114-54382

```

Query Match 96.6%; Score 3614; DB 15; Length 715;
Best Local Similarity 98.5%; Pred. No. 1.1e-190;
Matches 673; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

[illegible][illegible]

```
Db 334 GSSNQRETGGLSPSTPF-----GASNLLVNPLEPQ 364
QY 531 NADKIRVKIADIGNACWVHKHFTEDIQTRQYSIEVLIGAGYSTPADIWSTACMAFELAT 590
Db 365 NADKIKIKIADLGNACWVHKHFTEDIQTRQYRAVEVLIGAEYGPADIWSTACMAFELAT 424
QY 591 GDYLPFPHSGEDYSRDEDEHIAHIELLSIPRHFALSGKYSREFFNRRGELRHITKLPW 650
Db 425 GDYLPFPHSGEDYSRDEDEHIAHIELLSIPRHFALSGKYSREFFNRRGELRHITKLPW 484
QY 651 SLFDVLVEKYGWPHEDAAQFTDLIPMLEMVPKESASAGECLRHHPWLN 698
Db 485 GLYEVLMEXYEWPLEQATQFSAPLLPMMEYIPEKRSASAADCLQHPWLN 532

RESULT 9
US-10-723-860-1549
; Sequence 1549, Application US/10723860
; Publication No. US20040453606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1549
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-723-860-1549

Query Match 50.3%; Score 1882; DB 17; Length 533;
Best Local Similarity 56.5%; Pred. No. 1.5e-95;
Matches 366; Conservative 62; Mismatches 66; Indels 154; Gaps 7;

QY 52 TP-PEPEEILGSDDEEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLWCW 110
Db 38 TPVQMLQGLGSDDEEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLWCW 97
QY 111 DMQGRFVAMKVKSQAQHYTETALDEIKLLKCVRESDESDPNKDMVVLIDDFKISGNG 170
Db 98 DIQRKFVAMKVKSQAQHYTETALDEIKLLKCVRESDESDPNKDMVVLIDDFKISGNG 157
QY 171 IHVCMVFEVLGHLLKWIILKSNYQGLPVRCKVSIIRQVLQGLDYLHLSKCKIHTDIKPEN 230
Db 158 VHVMVFEVLGHLLKWIILKSNYQGLPVRCKVSIIRQVLQGLDYLHLSKCKIHTDIKPEN 217
QY 231 ILMCVDDAYVRMAAEATEWQKAGAPPPSGSAVSTAPQPKIGKISKNKKKLLKKQK 290
Db 218 ILLCVGDYAIRLAAEAATEWQKAGAPPPSGSAVSTAPQPKIGKISKNKKKLLKKQK 277
QY 291 AELLKRLQIEIELEAREAKIIEENITSAAPNDQGEYCPVEVKLTGTEEAARETA 350
Db 278 KRLLEERLRDQLRL-----AMEAATQAE-- 301
QY 351 KNGFAEDQEEKEDAEKENIEKEDDDVDQELANIDPTWIESPKTNGHTIENGPFSLQQLD 410
Db 302 -DSG----- 304
QY 411 DEDDDEDCPNPEEYNLDEPNAESDYTVSSYEQNGELPNGRHKIPESQPEPSTLSFS 470
Db 305 -----LRLDGGSG-----STSSGFGSL-----FSPASCSILS 333
QY 471 GSPLEPVACGSLSEGSPLTEQESSPHDRSRTVSASSTGDLPAKTRAADLLVNPDP 530
Db 334 GSSNQRETGGLSPSTPF-----GASNLLVNPLEPQ 364
```

```
QY 531 NADKIRVKIADIGNACWVHKHFTEDIQTRQYSIEVLIGAGYSTPADIWSTACMAFELAT 590
Db 365 NADKIKIKIADLGNACWVHKHFTEDIQTRQYRAVEVLIGAEYGPADIWSTACMAFELAT 424
QY 591 GDYLPFPHSGEDYSRDEDEHIAHIELLSIPRHFALSGKYSREFFNRRGELRHITKLPW 650
Db 425 GDYLPFPHSGEDYSRDEDEHIAHIELLSIPRHFALSGKYSREFFNRRGELRHITKLPW 484
QY 651 SLFDVLVEKYGWPHEDAAQFTDLIPMLEMVPKESASAGECLRHHPWLN 698
Db 485 GLYEVLMEXYEWPLEQATQFSAPLLPMMEYIPEKRSASAADCLQHPWLN 532

RESULT 10
US-10-369-493-5839
; Sequence 5839, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5839
; LENGTH: 698
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5839

Query Match 36.7%; Score 1374; DB 14; Length 698;
Best Local Similarity 46.4%; Pred. No. 1.8e-67;
Matches 304; Conservative 64; Mismatches 143; Indels 144; Gaps 13;

QY 53 PPEPEEILGSDDEEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLWCW 112
Db 96 PMDPGEQ-LGSDDEEDPADYCKGGYHPVIGDVFNARYHVIRKLGWGHFSTVWLWMT 154
QY 113 QGKRFVAMKVKSQAQHYTETALDEIKLLKCVRESDESDPNKDMVVLIDDFKISGNGIH 172
Db 155 QGKRFVAMKVKSQAQHYTETALDEIKLLKCVRESDESDPNKDMVVLIDDFKISGNGIH 214
QY 173 VCMVFEVLGHLLKWIILKSNYQGLPVRCKVSIIRQVLQGLDYLHLSKCKIHTDIKPENIL 232
Db 215 VAMVFEVLGCNLLKLIIRSNYRGLHGOVRKI CRQVLEALGYMHEKCGIHTDIKPENVL 274
QY 233 MCVDDAYVRMAAEATEWQKAGAPPPSGSAVSTAPQ---QKPIGKISKNKKKLLKKQK 289
Db 275 ITMSREEIKIWAQAAVAVKMMNM-KMGSSAVSTAPDHLVKAQENMTKNKKKKOMKKKAKK 333
QY 290 QAELEKRLQIEIELEAREAKIIEENITSAAPNDQGEYCPVEVKLTGTEEAAREAE 349
Db 334 QREKLEAEALAGLEI-----KMDANGLQAYNNAP 363
QY 350 AKDNGEAEQDE--KEDAEKENIEKEDDDVDQELANIDPTWIESPKTNGHTIENGPFSLQ 407
Db 364 ELENFNASQVEDVTMTETVNE-----GNRKVEIRSPDFDRTLTPFS--- 408
QY 408 QLDDEDDDEDCPNPEY--NLDEPNAESDYTVSSYEQNGELPNGRHKIPESQFPFS 465
Db 409 -----DPESKFGDLASPSAE---YLSS-----PMQOLP--- 433
QY 466 TSLFSGSLPVACGSLSEGSPLTEQESSPHDRSRTVSASSTGDLPAKTRAADLLVN 525
Db 434 ----PGGILPAP----- 441
```


QY 526 PLDPNRAD---KIRVKIADLGNACWVHKHFTEDIQTRYSLEVLIGAGYSTPADIWSTA 582
Db 442 PVPNIGDPCYCDLDVKIADLGNACWVHHYTTDDIQTRQRALEVLIGSGYPPADIWSTA 501
QY 583 CMAFELATGDLFEPHSGSDYSDRDEHIAHIELLGSIPRHFAISGYSKYSREFFNRGELR 642
Db 502 CMAFELATGDLFEPHSGSDYSDRDEHIAHIELLGAIPPSIYKKGKHWREFFHKNHLL 561
QY 643 HITKLPWSLFDVLVEKYGWPHEDAAQFTDFLIPMLEVPEKASAGECLRHPLW 697
Db 562 HIHQLKPSLYEVLRQKYESHEDAAQFESFLRPLMDFDQEKRSIAKALHPFL 616

RESULT 11
US-10-369-493-5838
; Sequence 5838, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2003-02-28
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 5838
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-5838

Query Match 36.5%; Score 1364.5; DB 14; Length 782;
Best Local Similarity 45.8%; Pred. No. 6.9e-67;
Matches 309; Conservative 75; Mismatches 174; Indels 117; Gaps 16;

QY 53 PPEEEILGSDDEQEDPADYCKGYHPVKIGDLFNGRYHVRKLGWGHFSTVWLCDWM 112
Db 96 PMPDGEQ-LGSDDEQEDPRDYKRGYHPVNIQGVFNARYHVRKLGWGHFSTVWLAWDT 154
QY 113 QOKRFVAMKVSAAHYTETALDEIKLLKCVRESPPDPKMDVVLQIDDFKISGMNGIH 172
Db 155 QOKRFVAMKVSAAHYTETALDEIKLLSVRSADENDIGCHKVQLLDFEFTVGTINGQH 214
QY 173 VCMVFEVLGHLHLLKWIISNYOGLPVRCKVSIIRQVLOGLDYLHSHCKIHTDIKPENIL 232
Db 215 VAMVFEVLGCLLKLIRSNYRGLHLEQVRKICQVLEALGYMHKCGIHTDIKPENVL 274
QY 233 MCVDDAYVRMAAEATEWOKACAPPPSGSAVSTAPQ---OKPIGKISXNKKKKLKKQKR 289
Db 275 ITMSREEIKIMAHAVARXNMN-KMSGSAVSTAPDHLVKMAQENWTKNKKKKKKRKK 333
QY 290 QAELEKRIQEIIELEEREAEKRIIEENITSAAPSDQDGEYCEPVKLTGTEEAEEAET 349
Db 334 QREKLEAEALGLEL-----KMDANGLOEA--YNN 361
QY 350 AKNDGEAEQOEKEDAEKENIEKDEDDVDQELANIDPTWIESPKTNGHIENGPFSLQOL 409
Db 362 APQNG-----VRMRPSSL-----LFNGP--IPQLL 384
QY 410 D-----DEDDDCPNPEYNLDEPNAESDYTYSSSYEQFN---GELPNGRHKIPESOF 461
Db 385 QSSSCVNTSPRTVPPLPPPLYPQVGCVCNQTYHLTVILNANVGELN----- 434
QY 462 PEFSTSLFSGSLPEVACGVLSEGSPLTEQESSSPSHDRSRTVS-----ASSTGDL--PK 514
Db 435 ---FNAS-----QVEDVTMETVNGNRRNKVRSRDRFRTTLTPSPDSEKFGDLASES 488

QY 515 AKTRAAD-----LLVNPDLPRNAD---KIRVKIADLGNACWVHKHFTEDIQTRQYR 562
Db 489 AYLSSPMSQLPPGGILPAPPVGNIGDPCYCDLDVKIADLGNACWVHHYTTDDIQTRQYR 548
QY 563 STEVLIGAGYSTPADIWSTACMAFELATGDLFEPHSGSDYSDRDEHIAHIELLGSIPR 622
Db 549 ALEVLIGSGYPPADIWSTACMAFELATGDLFEPHSGSDYSDRDEHIAHIELLGAIPP 608
QY 623 HPALSGKYSREFFNRGELRHITKLPWSLFDVLVEKYGWPHEDAAQFTDFLIPMLEMVP 692
Db 609 STYKKGKHWREFFHKNHLLHIHQLKPSLYEVLRQKYESHEDAAQFESFLRPLMDFDQ 668
QY 683 EKRSASAGECLRHPLW 697
Db 669 EKRSATAKIALHPFL 683

RESULT 12
US-10-425-114-49483
; Sequence 49483, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 49483
; LENGTH: 550
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700745847_FLI.pap
US-10-425-114-49483

Query Match 28.8%; Score 1076.5; DB 15; Length 550;
Best Local Similarity 39.2%; Pred. No. 3.2e-51;
Matches 255; Conservative 75; Mismatches 120; Indels 201; Gaps 19;

QY 63 SDDEQEDPADYCKGYHPVKIGDLFN-GRYHVRKLGWGHFSTVWLCDMGGKRFVAMK 121
Db 21 SEDEGTE---DYRRGGYHVRIGDAPFAGRYVYVQSGKLGWGHFSTVWLAWDTKHSRYALK 77
QY 122 VVKSQAQHYTETALDEIKLLKCVRESPPDPKMDVVLQIDDFKISGMNGIHVCMVFEVLG 181
Db 78 VQKSAQHYTEAANDITLIQIAEGDPPD--KCVVLLDHFHKGSPNGQHVCMVFEVLG 135
QY 182 HLLKWIISNYOGLPVRCKVSIIRQVLOGLDYLHSHCKIHTDIKPENILM--CVDVAY 239
Db 136 DNLLTLIKYSDVRGLPIAMVKEICFHLIAGLDYLHQQLSIHTDLKPENILLSTDD--- 192
QY 240 VRRMAAEATEWOKACAP---PPSGS--AVSTAPQOKP---IGKISNKKKKLKKKKQKRA 291
Db 193 -----PSKPRKSGAPLILPNSKDMAMESAGMDTKMLNGDLVKNHKKKKIKRKAQAA 246
QY 292 E-LLEKRIQEIIELEEREAEKRIIEENITSAAPSDQDGEYCEPVKLTGTEEAEEAETA 350
Db 247 HCVKEASE-----GVE----- 259
QY 351 KNDGEAEQOEKEDAEKENIEKDEDDVDQELANIDPTWIESPKTNGHIENGPFSLQQLD 410
Db 260 ---GNAE-----TSGAVES----- 270
QY 411 DEDDDEDCPNPEYNLDEPNAESDYTYSSSYEQFNGLPNGRHKIPESQPFESFSLPS 470

Db 271 -----SPNA-----SSAREQ-----TSSSA 285
QY 471 GSLEPVACGVSLS--EGSPLTEQESSPSHDSRSTVSASSTGDLPAKAKTRAADLLVNPDL 528
Db 286 GT-----SOLSDADGTYKLEQGNKGRSRWRQKLLAS-----VDL----- 320
QY 529 PRNADKIRKIADLGNACWVHKHFTEDIOIROYRSIEVLIGAGYSTPADIWSTACWAFEL 588
Db 321 -----KCKLVDFGNACWYKQFTNDIOIROYRCPREVILGSKYSTSADLWSFACICPEL 373
QY 589 ATGDVLFPHSGEDYSRDEDHIAHIIELGSGIPRHFALSGKYSRFFNRGELRHITK 648
Db 374 ATGDVLFPHSGDNFDRDEHIALMMLLGMWPRKIALGGYRSDFNRYGDLRHIRLR 433
QY 649 PWSLFDVLEKYGWPHEDAAQFTDFLIPMLEMVPKRSAGECLRHPLNS 699
Db 434 FWPLNKVLEKYLSEKDANDMTDFLVPILDFVPEKRTAGCQLLHPWNA 484

RESULT 13

US-10-424-599-215731
; Sequence 215731, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 215731
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_36833C.1.pep
US-10-424-599-215731

Query Match 28.6%; Score 1071.5; DB 15; Length 547;
Best Local Similarity 38.5%; Pred. No. 6.1e-51;
Matches 250; Conservative 76; Mismatches 126; Indels 197; Gaps 17;

QY 63 SDEQEDPADYCKGYHPVKIGDLFN-GRYHVIRKLGWGHFSTVWLCWDMQGRFVAMK 121
Db 18 SEDEGTE---DYRRGGYHVRIGDAFSAGRYVVOQSLGWHGFSTVWLAWDTKHSRYALK 74
QY 122 VVKSQAQHYTETALDEIKLLKCVRESPPDPNKMVVLIDDFKISGMNGIHVCMVFEVLG 181
Db 75 VQKSAQHYTEAAMDEITILQIAEGDPD--KKCVVKLLDHFKHSGPQGHVCMVFEVLG 132
QY 182 HLLKWIITKSNYQGLPVRCKSIIRQVLOGDYLHCKKIIHTDIKPENILM--CVDDAY 239
Db 133 DNLLTLIKYSDYRGVPLHMVKEICPHILVGLDYLHRELSIIHTDIKPENILLSTID--- 189
QY 240 VERMAAEATEWQKAGAPPPSSAVSTAPQKPIGK---AVSTAPQKPIGK---IGKNKKKKLKKKQK 291
Db 190 -----PSKDPKRSAGPLILPNSKQKMAWESGMDTKMLNGDLVQNHKKLIRKAKQAA 243
QY 292 E-LLEKRLQIEBELEREAEKIIIEENITSAFSDNDQGEYCPVKLKTGTEAEAEETA 350
Db 244 HCVKEKEASE-----GVE----- 256
QY 351 KONGEADEQEKEDAENIEKDEDDVDQELANIDPTWIESPKTNCHINGPFSLEQQLD 410
Db 257 ---GNAE-----TSGAVES----- 267
QY 411 DEDDDEEDCPNPEYNLDPENAESDYTYSSYEQFNGELPNGRHKIPESQPFESTSLFS 470
Db 268 -----SPNASSAREQASS-----SAGTSRLS 288

RESULT 14

US-10-739-930-5613
; Sequence 5613, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 5613
; LENGTH: 538
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-C1089_1.p
US-10-739-930-5613

Query Match 28.5%; Score 1067; DB 17; Length 538;
Best Local Similarity 37.2%; Pred. No. 1.1e-50;
Matches 242; Conservative 81; Mismatches 117; Indels 210; Gaps 16;

QY 63 SDEQEDPADYCKGYHPVKIGDLF-NGRYHVIRKLGWGHFSTVWLCWDMQGRFVAMK 121
Db 18 SEDEGTE---DYRRGGYHVRIGDTFRNGSYVIOQSLGWHGFSTVWLAWDTLNSRYALK 74
QY 122 VVKSQAQHYTETALDEIKLLKCVRESPPDPNKMVVLIDDFKISGMNGIHVCMVFEVLG 181
Db 75 IQKSAQHYTEAAMDEIKLQIAEGDAED--KKCVVKLLDHFKHAGPQGHVCMVFEVLG 132
QY 182 HLLKWIITKSNYQGLPVRCKSIIRQVLOGDYLHCKKIIHTDIKPENILM--VDDAY 239
Db 133 DNLLSVIKYSDYRGVPLHMVKEICPHILVGLDYLHRELSIIHTDIKPENILLCTIDP--- 190
QY 240 VERMAAEATEWQKAGAPPPSSAVSTAPQKPIGK---IGKNKKKKLKKKQK 288
Db 191 ---EADA---RKSGIPVLPTVKDQKAVPERPVEKEPKSYTSADLTNOKKKIRKK--- 241
QY 289 ROAELLEKRLQIEBELEREAEKIIIEENITSAFSDNDQGEYCPVKLKTGTEAEAEAE 348
Db 242 -----AKVSGSENERDS-----SNS- 259
QY 349 TAKONGEADEQEKEDAENIEKDEDDVDQELANIDPTWIESPKTNCHINGPFSLEQQLD 408
Db 260 -ARPNGNATVERLEBESSER---VKDAENVSO-----KSRGN----- 291
QY 409 LDEDDDEEDCPNPEYNLDPENAESDYTYSSYEQFNGELPNGRHKIPESQPFESTSL 468
Db 292 -----RRGSRSTRQK-----LLA----- 291
QY 469 FSGSLEPVACGVSLSGSPLETEQESSPSHDSRSTVSASSTGDLPAKAKTRAADLLVNPDL 528
Db 292 -----RRGSRSTRQK-----LLA----- 304

[illegible]

RESULT 15

```

US-10-369-493-12951
; Sequence 12951, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEIN
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 12951
; LENGTH: 561
; TYPE: PRT
; ORGANISM: Aspergillus nidulans
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(561)
; OTHER INFORMATION: unsure at all Xaa locations
US-10-369-493-12951

```

Query Match	28.2%;	Score 1054;	DB 14;	Length 561;
Best Local Similarity	37.6%;	Pred. No. 5.7e-50;		
Matches 247;	Conservative	85;	Mismatches 161;	Indels 164;
Gaps	13;			
Qy	57	EEETILGSDDEBOEPADYCKGGYHPVKIGDLF-NGRYHVIRKLGWGHFSTVWLCDWQGX	115	
Db	23	EEEE--ETTAEEESDYCKGGYHPVQIGELYNNGRYVVVRKLGWGHFSTVWLSRDTTG	80	
Qy	116	RFVAMKVKSQHYTETALDEIKLLCKCVRESDDPNKMDVYVQLIDDFKISGMNGIHVCM	175	
Db	81	KHVALKVVRSAHYTETAIDEIKLLNRIVQANPSHPGRKHVVSLDSFEHKGPGHVHVC	140	
Qy	176	VFEVLGHLLKWIITKSNVQGLPVRVCXSIIPVQLGLDYLHSCCKIHTHTDKPNNIMCV	235	
Db	141	VFEVLGNLLGLIKKNHNRGIPMLPVQKITQVLLGLDYLHRECGIHTHTDLKPNNLIE	200	
Qy	236	DDAYVRMAAEATWQKAGAPPPSGSAVSTAPQOKPIGKISKNKKKLLKKQKQAEILLE	295	
Db	201	GDV-----EQIVKTVYKKEAKKEQEDNNG-----	226	
Qy	296	KRLQIEELEREABRKIIENITGAAPSNDOGGYCPVKKLTGTGLEBAEAAETAKONGE	355	
Db	227	-----RRRRRTLTGSOPLSP-----LNTT-----PEFKHSSQNSH	258	
Qy	356	AEDOEKEDASKENIEDKDDVDQELANIDPTWIESPKTNCHIENGPSFLSQQLDDEDD	415	
Db	259	SSLQGVINESPCXNFQISENXT-----MLIGTGTSEAP-----SMQLLGIKDED	303	
Qy	416	EEDCPNPEYNLDBPNAESDYTSSSYEQFNGELPNGRHKIPESQFPFSTSLFSGSLEP	475	
Db	304	EOO-----KOREK-----TAYVER	317	

This Page Blank (uspio)

Page Blank (uspio)

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	3742	100.0	699	4	US-09-759-359A-2	Sequence 2, Appli
2	3742	100.0	699	4	US-10-207-973-2	Sequence 2, Appli
3	3620	96.7	698	2	US-09-016-000-2	Sequence 2, Appli
4	2070.5	55.3	655	1	US-08-364-002-2	Sequence 2, Appli
5	1374	36.7	1087	1	US-08-264-002-5	Sequence 5, Appli
6	979.5	26.2	544	1	US-08-364-002-7	Sequence 7, Appli
7	979.5	26.2	544	3	US-09-457-040B-15	Sequence 15, Appli
8	835	22.3	297	3	US-09-173-581-1	Sequence 1, Appli
9	835	22.3	297	3	US-09-420-915-1	Sequence 1, Appli
10	746.5	19.9	576	4	US-09-248-796A-20509	Sequence 20509, A
11	363	9.7	112	4	US-09-370-767-32637	Sequence 32637, A
12	363	9.7	112	4	US-09-270-767-47854	Sequence 47854, A
13	340.5	9.1	499	3	US-09-457-040B-13	Sequence 13, Appli
14	334.5	8.9	499	4	US-09-905-999-21	Sequence 21, Appli
15	321.5	8.6	490	4	US-09-376-594-314	Sequence 314, Appli
16	321.5	8.6	490	4	US-09-505-999-23	Sequence 23, Appli
17	318	8.5	568	2	US-08-835-170-4	Sequence 4, Appli
18	318	8.5	568	3	US-09-359-257-4	Sequence 4, Appli
19	318	8.5	588	2	US-08-835-170-2	Sequence 2, Appli
20	318	8.5	588	3	US-09-359-257-2	Sequence 2, Appli
21	317.5	8.5	429	4	US-09-810-671-5	Sequence 5, Appli
22	317.5	8.5	429	4	US-10-109-854-5	Sequence 5, Appli
23	317.5	8.5	429	4	US-10-339-656-5	Sequence 5, Appli
24	317.5	8.5	484	3	US-09-457-040B-12	Sequence 12, Appli
25	316	8.4	568	3	US-09-371-674-4	Sequence 4, Appli
26	316	8.4	588	3	US-09-371-674-2	Sequence 2, Appli
27	313.5	8.4	517	3	US-09-457-040B-14	Sequence 14, Appli

Db 361 EKDAEKENIEKDDDDVQDELANIPTWIESPKTNGHIENGPFSLFQQLDDDDDEEDCP 420
QY 421 NPEEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFSTSLFSGSLEPVACGS 480
Db 421 NPEEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFSTSLFSGSLEPVACGS 480
QY 481 VLSEGSPLTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNLDPNADKIRVKIA 540
Db 481 VLSEGSPLTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNLDPNADKIRVKIA 540
QY 541 DLGNACWVHKHFTEDIOTROYRSIEVLICAGYSTPADIWSTACMAFELATGDLPEPHSG 600
Db 541 DLGNACWVHKHFTEDIOTROYRSIEVLICAGYSTPADIWSTACMAFELATGDLPEPHSG 600
QY 601 EDYSRDEDHIAHIELLSGSI PRHFALSGKYSREFNRRGELRHITKLPWSLFDVLVEKY 660
Db 601 EDYSRDEDHIAHIELLSGSI PRHFALSGKYSREFNRRGELRHITKLPWSLFDVLVEKY 660
QY 661 GWPHEAQAQFTDFLIPMLEMVEPKRASAGECLRHPLWLS 699
Db 661 GWPHEAQAQFTDFLIPMLEMVEPKRASAGECLRHPLWLS 699

RESULT 2
US-10-207-973-2
; Sequence 2, Application US/10207973
; Patent No. 6753175
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROEF
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/10/207,973
; CURRENT FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 699
; TYPE: PRT
; ORGANISM: Human
US-10-207-973-2

Query Match. 100.0%; Score 3742; DB 4; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.7e-251;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRKVLAIQARKRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPEEI 60
Db 1 MSSRKVLAIQARKRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPPPPPPPEEI 60
QY 61 LQSDDEEDPADYCKGGVHPVKIGDLFNGRVHVIRKLGWGHFSTVWLCWDMQGRFVAM 120
Db 61 LQSDDEEDPADYCKGGVHPVKIGDLFNGRVHVIRKLGWGHFSTVWLCWDMQGRFVAM 120
QY 121 KVKSAQHYTETALDEIKLLKCVRESDFSPNKMVQVLIIDDFKISGNMGHIVCMVFEVL 180
Db 121 KVKSAQHYTETALDEIKLLKCVRESDFSPNKMVQVLIIDDFKISGNMGHIVCMVFEVL 180
QY 181 GHLLKWLKISYQGLPVCVKSIIRQVLOGLDYLHSCKIIHTDIKENTILMVCVDDAYV 240
Db 181 GHLLKWLKISYQGLPVCVKSIIRQVLOGLDYLHSCKIIHTDIKENTILMVCVDDAYV 240
QY 241 RMAAEATEWQAGAPPPSGSVAVSTAPQOKPIGKISNKKKLLKKKQKQAELEKRLQE 300
Db 241 RMAAEATEWQAGAPPPSGSVAVSTAPQOKPIGKISNKKKLLKKKQKQAELEKRLQE 300
QY 301 IEELEAEARKIIENITSAFNSDQDGEYCPEVKLKTTLGLEAAEAETAKDNGEAEQOE 360
Db 301 IEELEAEARKIIENITSAFNSDQDGEYCPEVKLKTTLGLEAAEAETAKDNGEAEQOE 360
QY 361 EKDAEKENIEKDDDDVQDELANIPTWIESPKTNGHIENGPFSLFQQLDDDDDEEDCP 420

Db 361 EKDAEKENIEKDDDDVQDELANIPTWIESPKTNGHIENGPFSLFQQLDDDDDEEDCP 420
QY 421 NPEEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFSTSLFSGSLEPVACGS 480
Db 421 NPEEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFSTSLFSGSLEPVACGS 480
QY 481 VLSEGSPLTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNLDPNADKIRVKIA 540
Db 481 VLSEGSPLTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNLDPNADKIRVKIA 540
QY 541 DLGNACWVHKHFTEDIOTROYRSIEVLICAGYSTPADIWSTACMAFELATGDLPEPHSG 600
Db 541 DLGNACWVHKHFTEDIOTROYRSIEVLICAGYSTPADIWSTACMAFELATGDLPEPHSG 600
QY 601 EDYSRDEDHIAHIELLSGSI PRHFALSGKYSREFNRRGELRHITKLPWSLFDVLVEKY 660
Db 601 EDYSRDEDHIAHIELLSGSI PRHFALSGKYSREFNRRGELRHITKLPWSLFDVLVEKY 660
QY 661 GWPHEAQAQFTDFLIPMLEMVEPKRASAGECLRHPLWLS 699
Db 661 GWPHEAQAQFTDFLIPMLEMVEPKRASAGECLRHPLWLS 699

RESULT 3
US-09-016-000-2
; Sequence 2, Application US/09016000
; Patent No. 5962232
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,000
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0465 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 688 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HEARNOT01
; CLONE: 307624


```
RESULT 5
US-08-264-002-5
; Sequence 5, Application US/08264002
; Patent No. 5559019
; GENERAL INFORMATION:
; APPLICANT: GUI, JIAN-FANG
; APPLICANT: FU, XIANG-DONG
; TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 Century Park East, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,002
; FILING DATE: 22-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3590
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1087 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; IMMEDIATE SOURCE:
; CLONE: CEHK
; NAME/KEY: Protein
; LOCATION: 1..1087
US-08-264-002-5

Query Match 36.7%; Score 1374; DB 1; Length 1087;
Best Local Similarity 46.7%; Pred. No. 5.6e-87;
Matches 305; Conservative 68; Mismatches 158; Indels 122; Gaps 13;

QY 53 PPEEEILGSDDEQEDPADYCKGYPVKGIDLPNGRYHVRKLGWGHSTVWLCDWM 112
DB 481 PNDPGEQ-LGSDDEQEDPRDYKRGYHPVNGDVFNARYHVRKLGWGHSTVWLAWDT 539
QY 113 QGKRVAMKVKSAGHYETALDEIKLLKCVRESPPDNKMDVYVQLIDDFKISGMNGIH 172
DB 540 QDKRVAMKVKSAGHYETALDEIKLLSVRSADPNIGCHVKVQLDFVTVGNGQH 599
QY 173 VCMVEVLGHLLKWKIISNVQGLPVRCKSIIRVQLGLDYLHKSCKIITHDIKPNIL 232
DB 600 VAMVPEVLGNLLKLIIRSNVRLGLEQVRKICRQVLEALGYMHEKCGIITHDIKPNVL 659
QY 233 MCVDDAYVRMAAEATEQKAGAPPPSGSAVSTAPQ---QKPIGKISNKKKLLKKKKR 289
DB 660 IIMSREEIKIMQAHAVARKMMN-KVSGSAVSTAPDHLVKMAQENWTKNKKKKKKKAKK 718
QY 290 QAELEKRLQETEELEBAERKIIIEENITSAAPSNDQGEYCEPVKLTKTGLEAAEAET 349
DB 719 QREKLEAELAGLEGLKMDANG--LQEAYNAPLTNIGKVMCNRRNGTLEENF---- 771
QY 350 AKNGEAEQBEKEPAEKENIEKEDDDVDQELANIDPTWIESPKTNNGHIENGPFSLQOL 409

Db 772 ---NASQVEDVTMEDTVNEN-----GNRNKVEIRSPDRFDRTRTLTPFS----- 811
QY 410 DDEDDDEDCNPPEY--NLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQFPEFS 467
DB 812 -----DPESKFGDLASPSAE---YLSS-----PMSQLP----- 836
QY 468 LFGSLPEFVACGVLSEGSPLTEQESPSHDSRTVSASSTGDLPKAKTRAADLLVNPL 527
DB 837 --PGGILPAP-----PV 846
QY 528 DPNRAD---KIRVKIADLGNACWVHKHFTEDIQTRYRSIEVLIGAGYSTPADINWSTACM 584
DB 847 GFNIGDPYCDIDVKIADLGNACWVNHHTDDIQTRYRALEVLIGSGYGPADINWSTACM 906
QY 585 AFELATGDLPEPHSGEDYSRDEDEHIAHIELGSIIPRHFAISGKYSREFFNRRELRI 644
DB 907 AFELATGDLPEPHSGEDYSRDEDEHLAHSSELLGAIPPSIYKKGKHWREFFHKNGHLLHI 966
QY 645 TKLKPWSLFDVLVEKYGWPHEDAQAQFTDPLPMLMVEPKEKASAGECLRHHPWL 697
DB 967 HQLKPWSLYEVLROKYEWSHEDAQQFESFLRPLDFDQEKSTAKIALKHPFL 1019

RESULT 6
US-08-264-002-7
; Sequence 7, Application US/08264002
; Patent No. 5559019
; GENERAL INFORMATION:
; APPLICANT: GUI, JIAN-FANG
; APPLICANT: FU, XIANG-DONG
; TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 Century Park East, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,002
; FILING DATE: 22-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3590
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 544 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-264-002-7

Query Match 26.2%; Score 979.5; DB 1; Length 544;
Best Local Similarity 36.6%; Pred. No. 5.5e-60;
Matches 233; Conservative 68; Mismatches 154; Indels 181; Gaps 12;

QY 68 QEDPADYCKGYPVKGIDLPNGRYHVI-RKLGWGHSTVWLCDWMQGRFVAMKVKS 126
DB 56 EENAEIDHYGYPVYIGEEFHRRYVVERKLGWGHSTVWLAYDBAAKRRVALKVRSA 115
QY 127 QHYTETALDEIKLLKCVRESPPDNKMDVYVQLIDDFKISGMNGIHVCMVEVLGHLLK 186
```



```
Db      116 EHYRETSIDEIRLOKIREGDEKHLGKHKHIIISLLDYFVHRGPNGAHVCVFEVLGENLLS 175
Qy      187 WIKSNYQGLPVCVKSIIRVOVLQGLDYLHSHCKIIHTDIPKBNILMVCVDDAYVRRMAAE 246
Db      176 LIQSYGHRGVPVGIKQIAYQLLIALDYLHRECGIIHTDLKPNVLCIDQDALQHEAP 235
Qy      247 ATEWOKAGAPPSPGSAVSTAPQOKPIGKISKNKKLKKQKROAELLEKRLQEIIEELER 306
Db      236 ATT-----SSPSTNTSSS-----KTRNNTGYTAKAPII-KRGQSVDNSAQ 274
Qy      307 EAERKIIENITSAAPSNQDGEYCEPVKLKTT-----GLEEAAEAETAKNGEADQEE 361
Db      275 --ERKTFAKNPTK---NSKPAGQVIPSPSTSTLSRFPSPLEGAVSISLRD----- 320
Qy      307 EAERKIIENITSAAPSNQDGEYCEPVKLKTT-----GLEEAAEAETAKNGEADQEE 361
Db      275 --ERKTFAKNPTK---NSKPAGQVIPSPSTSTLSRFPSPLEGAVSISLRD----- 320
Qy      362 KEDAENIEKEDDDVDQELANIDPTWIESPKTNHENGSPFSLFSGSLEPVACGVS 481
Db      321 -----SQHNSH- PNSSPFS-----SG 335
Qy      422 PSEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQFPPEFSTLSFSGSLEPVACGVS 481
Db      336 DNSLILDGVN-----GSQEPV----- 351
Qy      482 LSEGSLPTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNPDPNADKIRVKIAD 541
Db      352 -----PKITVKIAD 360
Qy      542 LGNACWVHKHFTEDIQTRYRSIEVLIGAGYSTPADINWSTACWAFELATGDLFEPHSGE 601
Db      361 LGNACWTRKHFTNDVQTRYRSPEVILGCRWGSADCSFACIIFELLTGDLFDPNGN 420
Qy      602 DYSRDEDHIAHIELLGSIPRHFALSGKYSRREFFNRGELRHITKLPKSLDFVLVEKYG 661
Db      421 SYSKEDDHIAQIIELLVNYPKQWALSGKHSRDLFNRGELRNHKLKFWPLKDVLEQKYH 480
Qy      662 WPHEDAAQFTDLIPMLEMVPKASAGCELRHPWL 697
Db      481 FSAELAQAQISDFLSPMLCFDPKRTNAGYMSNSPWL 516
```

RESULT 7

```
US-09-457-040B-15
; Sequence 15, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457, 040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Human
US-09-457-040B-15
```

```
Query Match      26.2%; Score 979.5; DB 3; Length 544;
Best Local Similarity 36.6%; Pred. No. 5.5e-60;
Matches 233; Conservative 68; Mismatches 154; Indels 181; Gaps 12;
```

```
Qy      68 QEDPADYKGGYHPVKIGDLFNGRYHVI-RKLGWGHFSTWLCWDMQGRFVAMKVKVSA 126
Db      56 EENAEDYHGGYHPVYIGEEFHHRYVVERKLGWGHFSTWLCWDMQGRFVAMKVKVSA 115
Qy      127 QHYTEALDEIKLLKCVRESDDSPDKMVMQVLIIDDFKISGNGNHHVCMVFEVLGHLLK 186
Db      116 EHYRETSIDEIRLOKIREGDEKHLGKHKHIIISLLDYFVHRGPNGAHVCVFEVLGENLLS 175
Qy      187 WIKSNYQGLPVCVKSIIRVOVLQGLDYLHSHCKIIHTDIPKBNILMVCVDDAYVRRMAAE 246
```

```
Db      176 LIQSYGHRGVPVGIKQIAYQLLIALDYLHRECGIIHTDLKPNVLCIDQDALQHEAP 235
Qy      247 ATEWOKAGAPPSPGSAVSTAPQOKPIGKISKNKKLKKQKROAELLEKRLQEIIEELER 306
Db      236 ATT-----SSPSTNTSSS-----KTRNNTGYTAKAPII-KRGQSVDNSAQ 274
Qy      307 EAERKIIENITSAAPSNQDGEYCEPVKLKTT-----GLEEAAEAETAKNGEADQEE 361
Db      275 --ERKTFAKNPTK---NSKPAGQVIPSPSTSTLSRFPSPLEGAVSISLRD----- 320
Qy      362 KEDAENIEKEDDDVDQELANIDPTWIESPKTNHENGSPFSLFSGSLEPVACGVS 481
Db      321 -----SQHNSH- PNSSPFS-----SG 335
Qy      422 PSEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQFPPEFSTLSFSGSLEPVACGVS 481
Db      336 DNSLILDGVN-----GSQEPV----- 351
Qy      482 LSEGSLPTEQESSPSHDSRTVSASSTGDLPAKTRAADLLVNPDPNADKIRVKIAD 541
Db      352 -----PKITVKIAD 360
Qy      542 LGNACWVHKHFTEDIQTRYRSIEVLIGAGYSTPADINWSTACWAFELATGDLFEPHSGE 601
Db      361 LGNACWTRKHFTNDVQTRYRSPEVILGCRWGSADCSFACIIFELLTGDLFDPNGN 420
Qy      602 DYSRDEDHIAHIELLGSIPRHFALSGKYSRREFFNRGELRHITKLPKSLDFVLVEKYG 661
Db      421 SYSKEDDHIAQIIELLVNYPKQWALSGKHSRDLFNRGELRNHKLKFWPLKDVLEQKYH 480
Qy      662 WPHEDAAQFTDLIPMLEMVPKASAGCELRHPWL 697
Db      481 FSAELAQAQISDFLSPMLCFDPKRTNAGYMSNSPWL 516
```

RESULT 8

```
US-09-173-581-1
; Sequence 1, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 119819
US-09-173-581-1
```

```
Query Match      22.3%; Score 835; DB 3; Length 297;
Best Local Similarity 45.1%; Pred. No. 2.7e-50;
Matches 188; Conservative 32; Mismatches 73; Indels 124; Gaps 10;
Qy      283 LKKQKQRAELLEKRLQEIIEELERAEARKIIEENITSAAPSNQDGEYCEPVKLKTTGLE 342
Db      1 MRKTKQKRLIEERLRLQRL----- 22
Qy      343 EAAEAETAKNGEADQEEKEKEDDAEKENTEKDEDDVDQELANIDPTWIESPKTNHENG 402
```

Db 23 EAMEAAT-----QAEDSGLRLDGGSGSTSSGCHP-----GGARAGP 59
Qy 403 FSLEQQLDDEDDDCPNPEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQFP 462
Db 60 -----SPASSAP-----CGGRSLSGSQT 81
Qy 463 EFTSLFSGSLPVPACGVLSEGSPLTQEESPSHDSRTVSAASSTGDL--PKAKTRAA 520
Db 82 GFGSLFS---PASC-SILSGSS---NQRE-----TGGLLSPSPFFGAS 118
Qy 521 DLLVNPDPNADKIRVKIADLGNACWVHKHFTEDIOTQRYRSIEVLIGAGYSTPADTWS 580
Db 119 NLLVNPDPNADKIRVKIADLGNACWVHKHFTEDIOTQRYRAVEVLIGAGYSTPADTWS 178
Qy 581 TACMAFELATGDLPEPHSGEDYSRDEDHIAHIELLGSIPRHFALSGKYSREFFNRGE 640
Db 179 TACMAFELATGDLPEPHSGEDYSRDEDHIAHIELLGSIPRHFALSGKYSREFFNRGE 238
Qy 641 LRHITKLKPSLFDVLVEKYGWPHEDAAQFTDFLIPMLEMVPKRSAGECLRHFWL 697
Db 239 LRHIHLKHGWLVEVLEMEKYEWPLEQATQFSALLPMNEYIPEKRASARDCLQHPWL 295

RESULT 9

US-09-420-915-1
; Sequence 1, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/420,915
; CURRENT FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/173,581
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 297
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 119819
US-09-420-915-1

Query Match 22.3%; Score 835; DB 3; Length 297;
Best Local Similarity 45.1%; Pred. No. 2.7e-50;
Matches 188; Conservative 32; Mismatches 73; Indels 124; Gaps 10;
Qy 283 LKKKQKQAELEKRLQELTEELEREAEARKIIEENITSAAPSNDQGEYCEVKLTGTLG 342
Db 1 MRKRKQKRLLEERLDRQL----- 22
Qy 343 EAAEAETAKDNGEAEQDEKEAEKENIEKEDVDQELANIDPTWIESPKTNHENGFP 402
Db 23 EAMEAAT-----QAEDSGLRLDGGSGSTSSGCHP-----GGARAGP 59
Qy 403 FSLEQQLDDEDDDCPNPEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIPESQFP 462
Db 60 -----SPASSAP-----CGGRSLSGSQT 81
Qy 463 EFTSLFSGSLPVPACGVLSEGSPLTQEESPSHDSRTVSAASSTGDL--PKAKTRAA 520
Db 82 GFGSLFS---PASC-SILSGSS---NQRE-----TGGLLSPSPFFGAS 118

Qy 521 DLLVNPDPNADKIRVKIADLGNACWVHKHFTEDIOTQRYRSIEVLIGAGYSTPADTWS 580
Db 119 NLLVNPDPNADKIRVKIADLGNACWVHKHFTEDIOTQRYRAVEVLIGAGYSTPADTWS 178
Qy 581 TACMAFELATGDLPEPHSGEDYSRDEDHIAHIELLGSIPRHFALSGKYSREFFNRGE 640
Db 179 TACMAFELATGDLPEPHSGEDYSRDEDHIAHIELLGSIPRHFALSGKYSREFFNRGE 238
Qy 641 LRHITKLKPSLFDVLVEKYGWPHEDAAQFTDFLIPMLEMVPKRSAGECLRHFWL 697
Db 239 LRHIHLKHGWLVEVLEMEKYEWPLEQATQFSALLPMNEYIPEKRASARDCLQHPWL 295

RESULT 10

US-09-248-796A-20509
; Sequence 20509, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20509
; LENGTH: 576
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20509

Query Match 19.9%; Score 746.5; DB 4; Length 576;
Best Local Similarity 31.3%; Pred. No. 8.6e-44;
Matches 188; Conservative 83; Mismatches 174; Indels 155; Gaps 13;
Qy 143 VRESPPDPNKDMVMVQLIDDFKISGMNGIHVCMVFEVLGHLLKWIKSNYQGLPVRCKV 202
Db 4 VTTSDIHHPGHQHVQLDFTTHKGPNGVHVVMVFEVLGENLLGLIRRYKRGIPVVFVK 63
Qy 203 SIIRQVQLGDLVHLSKCKIHTDIKPENILMCVDDAYVVRMAAEATEWQKAGAPPSGA 262
Db 64 QIAKQLLSALDFLRQCGVIHTDLKPENILIEIGD----- 98
Qy 263 VSTAPQKPIGKISKNKKKLKKQKQAELEKRLQELTEELEREAEARKIIEENITSAAP 322
Db 99 -----VEQIVKLVEE-ENLQKQLKRLSRTASKTSTP 129
Qy 323 -SNDQGEYCEVKLTGTLGEEAAEAETAKDN-----GEAB 357
Db 130 ISATPSSFSNHNANTTTTTTTAKKISINDSIISPTSSALTSSNFHNSPLSGRSG 189
Qy 358 DQEEKED--AKENIEKEDDDVDQELANI---DPTWIESPKTNHENGPPS----- 404
Db 190 RRTRRTTLITGSQPLPSPLRSFNKSFNTVYGLSSSTNTTPVRNISINSNNFINSNTT 249
Qy 405 -----LEQQLDDEDDDCPNPEYNLDEPNAESDYTYSSSYEQFNGELPNGRHKIP 457
Db 250 ANTTNGTATTTEDEEDGE-----NOTLNLSNMSITNSHN----- 288
Qy 458 ESQPEPSTSLFSGSLPVPACGVLSEGSPLTQEESPSHDSRTVSAASSTGDLPKAKT 517
Db 289 -----SNSYQPTATNV---DPVQIPQESNYSLDSS---SANVVED----- 323
Qy 518 RAADLLVNPDPNADKIRVKIADLGNACWVHKHFTEDIOTQRYRSIEVLIGAGYSTPAD 577
Db 324 -----IIN-----DNEILSVKIADLGNACWTHHFTEDIOTQRYRAPEILICYNGAUSD 373
Qy 578 TWSTACMAFELATGDLPEPHSGEDYSRDEDHIAHIELLGSIPRHFALSGKYSREFFNR 637

Db 374 LWSFACLIPELLTGDLFDPDRGKSYKDDHIAQIIEILGPPFNQMLKESYYAREFFNS 433
QY 638 RCELRHITKLPWSLFDVLVEKYGPHEDAAQFTFLIPMLEWPEKRSAGECRLHPWL 697
Db 434 RYELRRIMKLPWGLQDLVLEIKYKFLPNDIAIBSEFLFPLMKLPKEERADAGGMLNHPWL 493

RESULT 11

US-09-270-767-32637
; Sequence 32637, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32637
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32637

Query Match 9.7%; Score 363; DB 4; Length 112;
Best Local Similarity 54.7%; Pred. No. 4.3e-18;
Matches 70; Conservative 18; Mismatches 18; Indels 22; Gaps 2;
QY 488 LTEQESSPSHRSRTVSASSTGDLPAKAKTRAADLLVNPDPNADKIRVKIADLGNACW 547
Db 7 IPSQSQSSQNTYTIQS-----LIDN-----SNVRVKIADLGNACY 44
QY 548 VHKHFTEDIQROYRSIEVLIGAGYSTPADINSTACMAFELATGDLFPEPHSGEDYSRDE 607
Db 45 DVHHTFEDIQROYRSIEVLIGAGYSTPADINSTACMAFELATGDLFPEPHSGEDYSRDE 104
QY 608 DHIAHIE 615
Db 105 DHLAHIVD 112

RESULT 12

US-09-270-767-47854
; Sequence 47854, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47854
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47854

Query Match 9.7%; Score 363; DB 4; Length 112;
Best Local Similarity 54.7%; Pred. No. 4.3e-18;
Matches 70; Conservative 18; Mismatches 18; Indels 22; Gaps 2;
QY 488 LTEQESSPSHRSRTVSASSTGDLPAKAKTRAADLLVNPDPNADKIRVKIADLGNACW 547
Db 7 IPSQSQSSQNTYTIQS-----LIDN-----SNVRVKIADLGNACY 44
QY 548 VHKHFTEDIQROYRSIEVLIGAGYSTPADINSTACMAFELATGDLFPEPHSGEDYSRDE 607
Db 45 DVHHTFEDIQROYRSIEVLIGAGYSTPADINSTACMAFELATGDLFPEPHSGEDYSRDE 104
QY 608 DHIAHIE 615

Db 105 DHLAHIVD 112
||:||||:
US-09-457-040B-13
; Sequence 13, Application US/09457040B
; Patent No. 6387641
; GENERAL INFORMATION:
; APPLICANT: Vertex Pharmaceuticals Incorporated
; APPLICANT: Bellon, Steve
; TITLE OF INVENTION: Crystallized P38 Complexes
; FILE REFERENCE: VPI/98-14
; CURRENT APPLICATION NUMBER: US/09/457,040B
; CURRENT FILING DATE: 1999-12-08
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Human
US-09-457-040B-13
Query Match 9.1%; Score 340.5; DB 3; Length 499;
Best Local Similarity 18.6%; Pred. No. 1e-15;
Matches 121; Conservative 66; Mismatches 127; Indels 335; Gaps 13;
QY 69 EDPADYCKGGYHPVKIGDLFNGRYHVRIRKLGWGHFSTVWLCWD-MQGRFVAMKVKSQAQ 127
Db 144 EDDAE---GHLYHVGDWLOERYEIVSTLGBGTGFRVVQCVDRHRRGGARVALKIKNVE 199
QY 128 HYETALDEIKLLKCVRESDDPKMDVQVLIIDPFKISGMNGIHVCVFEVLGHLKWK 187
Db 200 KYKEAARLEINLVLEKINEKDP--NKNLCVQMFDFYHG-----HMCISFELLGLSTPDF 253
QY 188 IIKSNVQGLPVRCKVSIITROVLQGDYLSKCKIHTDIKPNILMCVDDAYVRRMAEA 247
Db 254 LKDNVLPYPIHQVRHMAFQLCQAVKFLHDN-KLTHTDLKPNILFV-----299
QY 248 TEMQKAGAPPPSGSAVSTAPOQKPIGKISKNKKLKKQKQKQKQKQKQKQKQKQKQKQKQ 307
Db 300 -----299
QY 308 AERKIIENITSAAPSNODGECYCEVKLKTGLEAAEAETAKONGEAEDQEKEDAOK 367
Db 300 -----299
QY 368 ENIEKDEDDVDOELANIDPTWIESPKTNNGHIENGPFSLQQLDDEDDDEDCPNPEYNL 427
Db 300 -----NSDYELT-----YNL 309
QY 428 DEPNAESDYTYSSSYEQFNGELPNGRHKIPESQPFESTLSFGSLSPVACGVLSEGPS 487
Db 310 EKKR-----313
QY 488 LTEQESSPSHRSRTVSASSTGDLPAKAKTRAADLLVNPDPNADKIRVKIADLGNACW 547
Db 314 -----DERSVKSTAVRVVDFGSATF 333
QY 548 VHKHFTEDIQROYRSIEVLIGAGYSTPADINSTACMAFELATGDLFPEPHSGEDYSRDE 607
Db 334 DHEHSTTVSTNTHYRAPEVILELWGSQPCDVMWSIGCIIFEYVYVFTLFTQTHD-----NR 387
QY 608 DHIAHIEILGSLP-----RHF-----ALSGKYSREFFNRRGELRHITKL 647
Db 388 EHLAMMERILGPIPSRMIRKTKQKYFYRGRLDWDENTSAGRYVRE--NCKPLRLYLTLS- 444
QY 648 KPSULFDVLVEKYGPHEDAAQFTFLIPMLEWPEKRSAGECRLHPW 696
Db 445 -----EABEHQFLDLESMLEYEPAKRLITLGEALQHPF 478

RESULT 14

Search completed: January 19, 2005, 15:42:20
Job time : 44 secs

Page Blank (uspio)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 19, 2005, 15:22:50 ; Search time 160 Seconds
(without alignments)
1567.199 Million cell updates/sec

Title: US-10-799-676-2
Perfect score: 3742
Sequence: 1 MSSRKVLAIQARKRPKREK.....MVPEKASAGELRHPWLNS 699

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq 23Sep04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3742	100.0	699	5	ABG32447 Human SR
2	3742	100.0	699	7	ADJ37689 Human kin
3	3626	96.9	688	6	Aao26722 SR protei
4	3626	96.9	688	8	ADJ96645 Human SRP
5	3622	96.8	675	6	Aao26721 SR protei
6	3620	96.7	688	2	AAY27053 Human pro
7	3586	95.8	686	5	Aau80373 Human cel
8	3586	95.8	686	6	Aae34823 Protein #
9	3586	95.8	686	8	ADOL9347 Human PRO
10	3395.5	90.7	681	7	Ades59481 Rat Prote
11	3117	83.3	1070	4	ABG02155 Novel hum
12	2070.5	55.3	655	2	Aaw06556 Serine ki
13	2050	54.8	654	6	Aao26720 SR protei
14	1890	50.5	567	8	Ade28302 Human KRP
15	1882	50.3	533	6	Aao26723 SR protei
16	1882	50.3	533	6	Aao26724 SR protei
17	1882	50.3	533	7	Ade31691 Human 142
18	1882	50.3	533	8	Adi29289 Human MAR
19	1882	50.3	533	8	ADQ18730 Human sof
20	1865.5	49.9	532	4	AAB65681 Novel pro
21	1338	35.8	338	6	Aao26725 SR protei
22	1337	35.7	334	8	Adn99237 Novel hum
23	1298.5	34.7	880	4	ABBY1719 Drosophil
24	1294.5	34.6	748	4	ABb62581 Drosophil
25	1075	28.7	538	3	AAG12830 Arabidops

26	1075	28.7	578	3	AAG12829	Aag12829 Arabidops
27	1068	28.5	538	3	AAG45141	Aag45141 Arabidops
28	1068	28.5	577	3	AAG45140	Aag45140 Arabidops
29	943	25.2	790	4	ABB63932	Abb63932 Drosophil
30	852	22.8	4134	2	AAY31946	Aay31946 Plasmodiu
31	841	22.5	452	3	AAG12831	Aag12831 Arabidops
32	835	22.3	297	3	AAY76748	Aay76748 Human pro
33	835	22.3	297	4	AAE06206	Aae06206 Human pro
34	835	22.3	297	5	ABB84433	Abb84433 Human pro
35	834	22.3	452	3	AAG45142	Aag45142 Arabidops
36	817	21.8	154	4	ABB38855	Abb38855 Peptide #
37	817	21.8	154	4	AAM32330	Aam32330 Peptide #
38	817	21.8	154	4	ABB23863	Abb23863 Protein #
39	817	21.8	154	4	AAM72065	Aam72065 Human bon
40	817	21.8	154	4	AAM59499	Aam59499 Human bra
41	817	21.8	154	4	ABG53749	Abg53749 Human liv
42	817	21.8	154	5	ABG41880	Abg41880 Human pep
43	720	19.2	187	4	AAB63273	Aab63273 Human bre
44	720	19.2	187	4	AAB63384	Aab63384 Human bre
45	666	17.8	386	3	AAG18143	Aag18143 Arabidops

ALIGNMENTS

RESULT 1
ABG32447
ID ABG32447 standard; protein; 699 AA.
XX
AC ABG32447;
XX
DT 20-NOV-2002 (first entry)
XX
DE Human SR protein-specific kinase 2, SRPK2.
XX
KW Human; kinase; SRPK2; SR protein-specific kinase 2; enzyme.
XX
OS Homo sapiens.
XX
PN US2002094560-A1.
XX
PD 18-JUL-2002.
XX
PF 16-JAN-2001; 2001US-00759359.
XX
PR 16-JAN-2001; 2001US-00759359.
XX
PA (ABUT/) ABU-THREIDEH J.
PA (GONG/) GONG F.
PA (KERC/) KETCHUM K A.
PA (DFRA/) DI FRANCESCO V.
PA (BEAS/) BEASLEY E M.
XX
PI Abu-Threideh J, Gong F, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-681805/73.
XX
DR N-PSDB; ABS52846.
XX
PT New SR protein-specific kinase 2 peptides and nucleic acid sequences,
XX
PT useful as models for developing human therapeutic targets, in identifying
XX
PT therapeutic proteins, and in identifying agents that modulate kinase
XX
PT activity.
XX
PS Claim 1; Fig 2; 56pp; English.
XX
CC The invention relates to an isolated human SR protein-specific kinase 2
XX
CC (SRPK2) peptide. The SR protein-specific kinase 2 peptide and nucleic
XX
CC acid sequences can be used as models for the development of human
XX
CC therapeutic targets, aid in the identification of therapeutic proteins,
XX
CC and as targets for the development of human therapeutic agents that
XX
CC modulate kinase activity in cells and tissues that express the kinase.
XX
CC These may further be used as query sequences to perform a search against
XX
CC sequence databases to identify other family members or related sequences.

CC The peptides can also be used to raise antibodies or to elicit another
CC immune response, as markers for tissues in which the corresponding
CC protein is preferentially expressed, to identify the binding
CC partner/ligand to develop a system to identify inhibitors of the binding
CC interaction, and in pharmacogenomic analysis. The nucleic acids are
CC useful as probes or primers, for expressing antigenic portions of the
CC proteins, for constructing vectors, host cells or transgenic animals
CC expressing the nucleic acids and peptide, for monitoring the
CC effectiveness of modulating compounds on the expression or activity of
CC the kinase gene in clinical trials or in treatment regimen, and as
CC antisense constructs to control kinase gene expression. The present
CC sequence represents the amino acid sequence of human SRPK2
XX
XX
SQ Sequence 699 AA;

Query Match 100.0%; Score 3742; DB 5; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1e-268;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRKVLAIQARKRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPEEEI 60
Db 1 MSSRKVLAIQARKRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPEEEI 60

QY 61 LGSDDDEEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCWDMQGRFVAM 120
Db 61 LGSDDDEEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCWDMQGRFVAM 120

QY 121 KVVKSQHYTETALDEIKLLKVCRESDDPNKDMVVLIDDFKISGMNGHVCWVFEVL 180
Db 121 KVVKSQHYTETALDEIKLLKVCRESDDPNKDMVVLIDDFKISGMNGHVCWVFEVL 180

QY 181 GHLLKWIISKSNYQGLPVRCKSIIRQVLQGLDYLHSCCKIIHTDIKPENILCMVDDAYV 240
Db 181 GHLLKWIISKSNYQGLPVRCKSIIRQVLQGLDYLHSCCKIIHTDIKPENILCMVDDAYV 240

QY 241 RMAAEATWQKAGAPPPSGSAVSTAPQOKPTGKISKNNKKLKKQKQAELEKRLQE 300
Db 241 RMAAEATWQKAGAPPPSGSAVSTAPQOKPTGKISKNNKKLKKQKQAELEKRLQE 300

QY 301 IEELEREARKIIEENITSAAPSNDDGYCEPVKLTGTLBEAAEAETAKNGEADQE 360
Db 301 IEELEREARKIIEENITSAAPSNDDGYCEPVKLTGTLBEAAEAETAKNGEADQE 360

QY 361 EKEDAEKENIEKDDVDQELANIDPTWIESPKTNHENGPPSLEQQDDEDDDEDCP 420
Db 361 EKEDAEKENIEKDDVDQELANIDPTWIESPKTNHENGPPSLEQQDDEDDDEDCP 420

QY 421 NPEEYNLDEPNASDYTSSSYQFNGELPNGRHKIPESQFPEFSTLSFGSLEPVACGS 480
Db 421 NPEEYNLDEPNASDYTSSSYQFNGELPNGRHKIPESQFPEFSTLSFGSLEPVACGS 480

QY 481 VLSEGSPLTEQESSPSHDSRTVSASTGDLFPKATRAADLIIVNPLDPNADKIRVKIA 540
Db 481 VLSEGSPLTEQESSPSHDSRTVSASTGDLFPKATRAADLIIVNPLDPNADKIRVKIA 540

QY 541 DLGNACWVKHFTEDIQYRSIEVLIGAGYSTPADIWSTACMAFELATGDFLPEPHSG 600
Db 541 DLGNACWVKHFTEDIQYRSIEVLIGAGYSTPADIWSTACMAFELATGDFLPEPHSG 600

QY 601 EDYSRDEDHIAHIELLGSIPRFALSGKYSREFFNRRGELRHITKLPWSLFDVLVEKY 660
Db 601 EDYSRDEDHIAHIELLGSIPRFALSGKYSREFFNRRGELRHITKLPWSLFDVLVEKY 660

QY 661 GWPHEDAQAQTFDILPMLNVEPKRASAGECLRHFWLNS 699
Db 661 GWPHEDAQAQTFDILPMLNVEPKRASAGECLRHFWLNS 699

RESULT 2
ADJ37689
ID ADJ37689 standard; protein; 699 AA.
XX
AC ADJ37689;

XX 06-MAY-2004 (first entry)
DT Human kinase protein.
DE
DE Kinase protein; drug screening assay; gene therapy;
KW pharmacogenomic analysis; testis; nervous tissue; fetal; lung;
KW brain anaplastic oligodendroglioma; lung carcinoma tissue;
KW soft tissue leiomyosarcoma; ovary tumour tissue; germ cell tumour tissue;
KW transgenic animal; human; enzyme.
XX Homo sapiens.
XX US2003175927-A1.
FN 18-SEP-2003.
XX 31-JUL-2002; 2002US-00207973.
PF 16-JAN-2001; 2001US-00759359.
PR (APPL-) APPLERA CORP.
XX Abu-Threideh J, Gong F, Ketchum KA, Di Francesco V, Beasley EM;
PI WPI; 2003-898544/82.
XX N-PSDB; ADJ37688.
DR New peptides related to kinase protein subfamily useful for treating
XX disorders associated with abnormal expression of kinase protein in
XX testis, nervous tissue, fetal, lung, ovary tumor tissue.
XX Claim 1; SEQ ID NO 2; 102pp; English.
PS The present invention relates to kinase protein and nucleotides encoding
CC them. The invention is useful in drug screening assays, tissue typing,
CC gene therapy and pharmacogenomic analysis. They are also useful in
CC treating disorders associated with the absence of, inappropriate, or
CC unwanted expression of kinase protein in testis, nervous tissue, fetal,
CC lung, brain anaplastic oligodendroglioma, lung carcinoma tissue, soft
CC tissue leiomyosarcoma, ovary tumour tissue, or germ cell tumour tissue.
CC The invention are useful as models for the development of human
CC therapeutic targets, aid in the identification of therapeutic proteins
CC and serve as targets for the development of human therapeutic agents that
CC modulate protease activity in cells and tissues that express the kinase
CC peptide. The invention is also useful in producing a kinase protein or
CC peptide, and non-human transgenic animals. The present sequence is human
CC kinase protein.
XX
SQ Sequence 699 AA;

Query Match 100.0%; Score 3742; DB 7; Length 699;
Best Local Similarity 100.0%; Pred. No. 1.1e-268;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSSRKVLAIQARKRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPEEEI 60
Db 1 MSSRKVLAIQARKRPKREKHPKPEQOKAPLVPPPPPPPPPPPPPPPPPPPPPEEEI 60

QY 61 LGSDDDEEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCWDMQGRFVAM 120
Db 61 LGSDDDEEQEDPADYCKGGYHPVKIGDLFNGRYHVIRKLGWGHFSTVWLCWDMQGRFVAM 120

QY 121 KVVKSQHYTETALDEIKLLKVCRESDDPNKDMVVLIDDFKISGMNGHVCWVFEVL 180
Db 121 KVVKSQHYTETALDEIKLLKVCRESDDPNKDMVVLIDDFKISGMNGHVCWVFEVL 180

QY 181 GHLLKWIISKSNYQGLPVRCKSIIRQVLQGLDYLHSCCKIIHTDIKPENILCMVDDAYV 240
Db 181 GHLLKWIISKSNYQGLPVRCKSIIRQVLQGLDYLHSCCKIIHTDIKPENILCMVDDAYV 240

QY 241 RMAAEATWQKAGAPPPSGSAVSTAPQOKPTGKISKNNKKLKKQKQAELEKRLQE 300
Db 241 RMAAEATWQKAGAPPPSGSAVSTAPQOKPTGKISKNNKKLKKQKQAELEKRLQE 300

Db 241 RRMAEATEWQKAGAPPPGSAVSTAPQOKPIGKISKNNKKLKKQKRAELLEKRLQE 300
Qy 301 IBELEAREARKIIEENITSAAPNSNDQGEYCPVVKLTGTLGEAAAEATAKONGEADQOE 360
Db 301 IBELEAREARKIIEENITSAAPNSNDQGEYCPVVKLTGTLGEAAAEATAKONGEADQOE 360
Qy 361 EKEDAENIEKDEDDVDOELANIDPTWIESPKTNHENGPPSLQQLDDEDDDEDCP 420
Db 361 EKEDAENIEKDEDDVDOELANIDPTWIESPKTNHENGPPSLQQLDDEDDDEDCP 420
Qy 421 NPEEYNLDPEPNAESDYTSYSSYEQFNGELPNGRHKIPESQFPFSTSLFSGSLPEPVACGS 480
Db 421 NPEEYNLDPEPNAESDYTSYSSYEQFNGELPNGRHKIPESQFPFSTSLFSGSLPEPVACGS 480
Qy 481 VLSEGSPLTEQEBSSPSHRSRTVSASSTGDLPAKATRAADLLVNPDPNADKIRVKIA 540
Db 481 VLSEGSPLTEQEBSSPSHRSRTVSASSTGDLPAKATRAADLLVNPDPNADKIRVKIA 540
Qy 541 DLGNACVWHKHFTEDIQTRQYRSIEVLIGAGYSTPADINWSTACMAFELATGDLFEPHSG 600
Db 541 DLGNACVWHKHFTEDIQTRQYRSIEVLIGAGYSTPADINWSTACMAFELATGDLFEPHSG 600
Qy 601 EDYSRDEDHIAHIELLGSIPRHFALSGKYSREFFNRRGELRHITKLPWSLFDVLVEKY 660
Db 601 EDYSRDEDHIAHIELLGSIPRHFALSGKYSREFFNRRGELRHITKLPWSLFDVLVEKY 660
Qy 661 GWPHEDAAOFTDPLIMLEWPEKRASAGECLRHHPWLS 699
Db 661 GWPHEDAAOFTDPLIMLEWPEKRASAGECLRHHPWLS 699

RESULT 3

AA026722
ID AA026722 standard; protein; 688 AA.

XX AAO26722;

XX 27-MAR-2003 (first entry)

DE SR protein-specific kinase-1 protein, SEQ ID No 15.

XX Cytostatic; p53 pathway-modulating agent; SR protein; cancer; apoptosis;
KW lung cancer; breast cancer; serine/arginine protein-specific kinase;
KW colon cancer; ovarian cancer; SRPK; cell proliferation; human.

XX Homo sapiens.

XX WO200299427-A1.

XX 12-DEC-2002.

XX 03-JUN-2002; 2002WO-US017525.

XX 05-JUN-2001; 2001US-0296076P.

XX 10-OCT-2001; 2001US-0328605P.

XX 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.

XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;

XX WPI; 2003-156865/15.

XX Identifying candidate p53 pathway-modulating agents useful as therapeutic
PT targets for disorders related with defective p53 function, by screening
PT for agents modulating serine/arginine protein-specific kinase activity.
XX Claim 13; Page 127-130; 137pp; English.

PS The invention relates to a novel method for identifying candidate p53

XX pathway-modulating agents. The novel method comprises screening for
CC agents that modulate the activity of a serine/arginine protein-specific
CC kinase. The invention provides methods for utilising p53 modifier genes

CC and polypeptides to identify candidate therapeutic agents that can be
CC used in the treatment of disorders associated with defective p53
CC function. The methods are also useful for modulating a p53 pathway in a
CC mammalian cell, or for diagnosing or treating a disease associated with
CC defective p53 function, e.g. cancers such as breast cancer, colon cancer,
CC lung cancer or ovarian cancer. The serine/arginine protein-specific
CC kinase (SRPK) polypeptides and nucleic acids are useful for identifying
CC and testing agents that modulate SRPK function. The animal models are
CC useful for in vivo assays to test the activity of a candidate p53-
CC modulating agent, or to assess the role of SRPK in a p53 pathway process
CC such as apoptosis or cell proliferation. This sequence represents a
CC protein of a serine/arginine-rich domain protein-specific kinase-1 of the
CC invention
XX

XX Sequence 688 AA;

Query Match 96.9%; Score 3626; DB 6; Length 688;

Best Local Similarity 98.8%; Pred. No. 4.3e-260;

Matches 675; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 17 KREKHPKKPEPOQKAPLVPPPPPPPPPPPPPPDPPTPEPEBEILGSDDEEQBDPADYCK 76
Db 6 EKSSSSERPEPOQKAPLVPPPPPPPPPPPPPPDPPTPEPEBEILGSDDEEQBDPADYCK 65
Qy 77 GGYHPVKIGDLFNGRHYVIRKLGWGHFSTVWLCWDMQGRFVAMKVVKSAQHYTETALDE 136
Db 66 GGYHPVKIGDLFNGRHYVIRKLGWGHFSTVWLCWDMQGRFVAMKVVKSAQHYTETALDE 125
Qy 137 IKLLKCVRESDPDPNKMVVQIIDDFKISGMNGIHVCMVFEVLGHLLKWIISNYQGL 196
Db 126 IKLLKCVRESDPDPNKMVVQIIDDFKISGMNGIHVCMVFEVLGHLLKWIISNYQGL 185
Qy 197 PVRCVKSIIRQVLQGLDYLHSCCKIHTDKPENILMVDVDAVVRMAAEATEWQKAGAP 256
Db 186 PVRCVKSIIRQVLQGLDYLHSCCKIHTDKPENILMVDVDAVVRMAAEATEWQKAGAP 245
Qy 257 PPSGSAVSTAPQOKPIGKISKNNKKLKKQKRAELLEKRLQEIELEAREARKIIEEN 316
Db 246 PPSGSAVSTAPQOKPIGKISKNNKKLKKQKRAELLEKRLQEIELEAREARKIIEEN 305
Qy 317 ITSAPNSNDQGEYCPVVKLTGTLGEAAAEATAKONGEADQOEKEKENIEKDEDD 376
Db 306 ITSAPNSNDQGEYCPVVKLTGTLGEAAAEATAKONGEADQOEKEKENIEKDEDD 365
Qy 377 VDOELANIDPTWIESPKTNHENGPPSLQQLDDEDDDEDCPNPEEYNLDPEPNAESDY 436
Db 366 VDOELANIDPTWIESPKTNHENGPPSLQQLDDEDDDEDCPNPEEYNLDPEPNAESDY 425
Qy 437 TYSSSYEQFNGELPNGRHKIPESQFPFSTSLFSGSLPEPVACGSVLSEGSPLTEQEBSSP 496
Db 426 TYSSSYEQFNGELPNGRHKIPESQFPFSTSLFSGSLPEPVACGSVLSEGSPLTEQEBSSP 485
Qy 497 SHDRSRTVSASSTGDLPAKATRAADLLVNPDPNADKIRVKIADLGNACVWHKHFTEDI 556
Db 486 SHDRSRTVSASSTGDLPAKATRAADLLVNPDPNADKIRVKIADLGNACVWHKHFTEDI 545
Qy 557 QTRQYRSIEVLIGAGYSTPADINWSTACMAFELATGDLFEPHSGEDYSRDEDHIAHIEL 616
Db 546 QTRQYRSIEVLIGAGYSTPADINWSTACMAFELATGDLFEPHSGEDYSRDEDHIAHIEL 605
Qy 617 LGSIPRHFALSGKYSREFFNRRGELRHITKLPWSLFDVLVEKYGHPHEDAAOFTDPLIP 676
Db 606 LGSIPRHFALSGKYSREFFNRRGELRHITKLPWSLFDVLVEKYGHPHEDAAOFTDPLIP 665
Qy 677 MLEWPEKRASAGECLRHHPWLS 699
Db 666 MLEWPEKRASAGECLRHHPWLS 688

RESULT 4

ADJ96645

ID ADJ96645 standard; protein; 688 AA.

XX

Dd 605 LGSIPRHFALSKYSEFNRNGERLHITKLKPSLFDFVLVEKYGNWPHEDAAQFTDPLP 684

Qy 677 MLEMWPEKRASAGECRLHPWLNS 699
|||||||
Db 665 MLEMWPEKRASAGEC-RHPWLNS 686

RESULT 10
ADES59481
ID ADE59481 standard; protein; 681 AA.
XX ADE59481;
XX 29-JAN-2004 (first entry)
XX Rat Protein NP_033300, SEQ ID NO 5376.
DX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
XW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX Rattus norvegicus.
OS WO2003016475-A2.
PN PD 27-FEB-2003.
XX 14-AUG-2002; 2002WO-USO25765.
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-034382P.
PR 26-NOV-2001; 2001US-0333347P.
XX (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; NP_033300.
XX New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

Pt Claim 1; Page; 1017pp; English.

Cc The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. Gene Therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Cc XX

XX Chien D, Jin P, Hawkins PR, Baughn MR, Becha SD, Chang H;
PI Ding L, Elliott VS, Emerling BM, Gandhi AR, Gietzen KJ, Griffin JH;
PI Gururajan R, Hafalia AJA, Ison CH, Kable AE, Khare R, Lee SY;
PI Lee EA, Lu Y, Marquis JP, Lehr-Mason PM, Ramkumar J, Richardson TW;
PI Swarnakar A, Tran UK, Chawla NK, Yao MG, Yue H, Bhatia U;
PI Burrill JD, Lee S, Blake JJ, Ho A, Zheng W;
XX WPI: 2004-011523/01.
DR N-PSDB; ADE28354.
XX
XX New human kinases and phosphatases, and polynucleotides encoding them,
PT useful for treating, preventing or diagnosing e.g. cell proliferative
PT disorders, inflammatory, autoimmune, viral, bacterial, parasitic or
PT fungal diseases.
XX
PS Claim 1; SEQ ID NO 13; 340pp; English.
XX
XX The invention relates to a novel isolated kinase and phosphatase (KPP)
CC polypeptide. The polypeptide of the invention demonstrates hepatotropic,
CC antiarteriosclerotic, antiposoriatic, cytostatic, haemostatic, muscular,
CC cerebroprotective, nootropic, ophthalmological, anticonvulsant,
CC vasotropic, neuroprotective, antiparkinsonian, antiasthmatic,
CC antianemic, antiasthmatic, antidiabetic, antiinflammatory, osteopathic,
CC antiarthritic, antirheumatic, dermatological, virucide, antibacterial,
CC fungicide, antiparasitic, protozoicidal, antihelminthic, antitumor,
CC cardiovascular, antiarteriosclerotic and immunosuppressive activities.
CC The KPP polypeptides may be useful for diagnosing, treating or preventing
CC cell proliferative disorders including cirrhosis, hepatitis,
CC arteriosclerosis, psoriasis, primary thrombocytopenia and cancer,
CC developmental disorders such as renal tubular acidosis, Becker's muscular
CC dystrophy, gonadal dysgenesis, hypothyroidism or seizures, neurological
CC disorders e.g. Pick's disease, cataract, epilepsy, ischaemic
CC cerebrovascular disease, stroke, Alzheimer's disease, Parkinson's disease
CC or dementia, autoimmune or inflammatory disorders including AIDS,
CC allergies, anaemia, asthma, diabetes mellitus, bronchitis, osteoporosis,
CC osteoarthritis, rheumatoid arthritis, contact dermatitis or gout and
CC lipid disorders such as cholestasia, Gaucher's disease, diabetes,
CC atherosclerosis or liver disease, as well as viral, bacterial, fungal,
CC parasitic, protozoan or helminthic infections and trauma. Furthermore,
CC the polypeptide may be utilised during gene therapy procedures. The
CC current sequence is that of the human KPP protein of the invention.
XX
SQ Sequence 567 AA;
Query Match 50.5%; Score 1890; DB 8; Length 567;
Best Local Similarity 58.1%; Pred. No. 2.2e-131;
Matches 378; Conservative 57; Mismatches 90; Indels 126; Gaps 12;
QY 52 TP-PEPEEILGSDDEQEDPADYCKGGYHPVKIGDLNFRVYVRKLGWGHFSTWLCW 110
DB 38 TPVPQMLQGLGSDDEQEDPADYCKGGYHPVKIGDLNFRVYVRKLGWGHFSTWLCW 97
QY 111 DMQGRFVAMKVKSAOHVYETALDEIKLLKCVRESPPDKMVMVQLIDPFKISGNG 170
DB 98 DIQRKRFVAMKVKSAOHVYETALDEIKLLKCVRESPPDKMVMVQLIDPFKISGNG 157
QY 171 IHVCMVFYGLHLLKWIISNYQGLPVRCKVSIIRQVLQGLDYLSKCKIHTDKPEN 230
DB 158 VHVCVMVLEVLGHLKWIISNYQGLPVRCKVSIIRQVLQGLDYLSKCKIHTDKPEN 217
QY 231 ILMCVDDAYVRMAEATEWQAGAPPSPGSAVSTAPQO-KPIGKISNKKKLLKKOKR 289
DB 218 ILLCVGDYVIRLAAEATEWQAGAPPSPGSAVSTAPQOVLQTKLSNKKRKKRKRQ 277
QY 290 QAELEKRLQETEEELEREAERKIIENITSAPSNDQGEYCEPVKLTGTGLEEAET 349
DB 278 QKRLLEERLDRQLR-----EAEARAT 299
QY 350 AKDNGEAEQEEKEAEKENIEKEDDDVDQELANIDPTWIESPKTNGHIENGPFSLQOL 409
DB 300 -----QAEGLSLDGGSGSTSSGCHP-----GGARAGP----- 329

QY 410 DDEDDDEEDCPNPEYNLDPENAESDYTYSSSYEQFNGELPNGRHKIPESQFPERSTSLF 469
DB 330 -----SPASSSPAP-----GGRSLSAGSQTSGRSGSLF 358
QY 470 SGLSEPVACGVLSESGSLTEQEESSPSHRSRTVTSASSTGDL--PKAKTRAADLLVNPL 527
DB 359 S-----PASC-SILSGSS--NQRE-----TGGLSPSTPFGASNLLVNPL 395
QY 528 DPNADKIRVKIADLGNACWVHKHFTEDIQTRQYRSIEVLICAGYSTPADISTACMAFE 587
DB 396 EPQNAADKIKIADLGNACWVHKHFTEDIQTRQYRAVEVLICAGYGGPPADISTACMAFE 455
QY 588 LATGDLPEPHSGEDYSRDEDHIAHIELGSIIPHFALSGKYSREFFNRRGELRHITKL 647
DB 456 LATGDLPEPHSGEDYSRDEDHIAHIELGDIPPAPALSGKYSREFFNRRGELRHITKL 515
QY 648 KPWSLFDVLVEKYGWPHEDAQAQTFDLIPMLEMVPKRSASAGECLRHHPWLN 698
DB 516 KHWGLYEVLMKEYEWPLEQATQPSAFLPMPMEYIPEKRSAAADCLQHPWLN 566
RESULT 15
ID AAO26723
XX AAO26723 standard; protein; 533 AA.
XX AAO26723;
XX
DT 27-MAR-2003 (first entry)
XX
DE SR protein-specific kinase-1 protein, SEQ ID NO 16.
XX
KW Cytostatic; p53 pathway-modulating agent; SR protein; cancer; apoptosis;
KW lung cancer; breast cancer; serine/arginine protein-specific kinase;
KW colon cancer; ovarian cancer; SRPK; cell proliferation; human.
OS Homo sapiens.
PN WO2002399427-A1.
XX
XX 12-DEC-2002.
XX
XX 03-JUN-2002; 2002WO-US017525.
XX
XX 05-JUN-2001; 2001US-0296076P.
XX 10-OCT-2001; 2001US-0328605P.
XX 15-FEB-2002; 2002US-0357253P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
WPI; 2003-156865/15.
XX
XX Identifying candidate p53 pathway-modulating agents useful as therapeutic
XX targets for disorders related with defective p53 function, by screening
XX for agents modulating serine/arginine protein-specific kinase activity.
PS Claim 13; Page 130-132; 137pp; English.
XX
XX The invention relates to a novel method for identifying candidate p53
XX pathway-modulating agents. The novel method comprises screening for
XX agents that modulate the activity of a serine/arginine protein-specific
XX kinase. The invention provides methods for utilising p53 modifier genes
XX and polypeptides to identify candidate therapeutic agents that can be
XX used in the treatment of disorders associated with defective p53
XX function. The methods are also useful for modulating a p53 pathway in a
XX mammalian cell, or for diagnosing or treating a disease associated with
XX defective p53 function, e.g. cancers such as breast cancer, colon cancer,
XX lung cancer or ovarian cancer. The serine/arginine protein-specific
XX kinase (SRPK) polypeptides and nucleic acids are useful for identifying
XX and testing agents that modulate SRPK function. The animal models are
XX useful for in vivo assays to test the activity of a candidate p53-
XX modulating agent, or to assess the role of SRPK in a p53 pathway process


```
QY 877 TTATGGAGAACCGCTCGACGAGATAGAAATTTGGAGCGAGAGCTGAAAGGAAATA 936
Db 61 TTATTTGGAGAACCGCTCGAGAGATAGAAATTTGGAGCGAGAGCTGAAAGGAAATA 120
QY 937 ATAGAAGAAAACATCACCTCAGCTGCACTTCCAAATGACAGGATGGCGAAATCTGCCCCA 996
Db 121 ATAGAAGAAAACATCACCTCAGCTGCACTTCCAAATGACAGGATGGCGAAATCTGCCCCA 180
QY 997 GAGGTGAACCTTAAAAACAAACAGATTAGAGAGGGCGCTGAGGACGACATGCAAGGAC 1056
Db 181 GAGGTGAACCTTAAAAACAAACAGATTAGAGAGGGCGCTGAGGACGACATGCAAGGAC 240
QY 1057 AATGCTGAGCTGAGGACGAGGACGAGAAAGAGATGCTGAGAAAGAAAACATTTGAAAAA 1116
Db 241 AATGCTGAGCTGAGGACGAGGACGAGAAAGAGATGCTGAGAAAGAAAACATTTGAAAAA 300
QY 1117 GATGAAGATGATGTAGATCAGAACTTTCGAAATAGACCCCTACCTGATGAAATCACCT 1176
Db 301 GATGAAGATGATGTAGATCAGAACTTTCGAAATAGACCCCTACCTGATGAAATCACCT 360
QY 1177 AAAACCAATGGCCATATTGAGATGGCCCATTTCTCACTGGAGCAGCAACTGGACGATGAA 1236
Db 361 AAAACCAATGGCCATATTGAGATGGCCCATTTCTCACTGGAGCAGCAACTGGACGATGAA 420
QY 1237 GATGATGATGAAGAGACTGCCAAATCTGAGGAATATAATTTGATGAGCCAAATGCA 1296
Db 421 GATGATGATGAAGAGACTGCCAAATCTGAGGAATATAATTTGATGAGCCAAATGCA 480
QY 1297 GAAAGTGATTACACATATAGCAGCTCTATGAACAATTCATGTTGAAATGCAATGGA 1356
Db 481 GAAAGTGATTACACATATAGCAGCTCTATGAACAATTCATGTTGAAATGCAATGGA 540
QY 1357 CGACATAAAATTTCCGAGTACAGTTCCAGAGTTTTCACCTGCTGTTCTCTGATGCC 1416
Db 541 CGACATAAAATTTCCGAGTACAGTTTCCAGAGTTTTCACCTGCTGTTCTCTGATGCC 600
QY 1417 TTGAACCTTGCGCTGCGCTCTGTGCTTCTGAGGGATCACCACTTACTGAGCAAGAG 1476
Db 601 TTGAACCTTGCGCTGCGCTCTGTGCTTCTGAGGGATCACCACTTACTGAGCAAGAG 660
QY 1477 GAGAGCAGTCCATCCATGACAGAGAGAAACGGTTTCAGCTCAGCTAGTGGGATTTG 1536
Db 661 GAGAGCAGTCCATCCATGACAGAGAGAAACGGTTTTCAGCTCAGCTAGTGGGATTTG 720
QY 1537 CCAAGAGCAAAACCCGGCGAGCTGCTGTTGTTGTAATCCCTGATCCGCGGAATGCA 1596
Db 721 CCAAGAGCAAAACCCGGCGAGCTGCTGTTGTTGTAATCCCTGATCCGCGGAATGCA 780
QY 1597 GATAAAATTAGAGTAAAAATTTGCTGACCTGGGAAATGCTTGTGTTGGTGATATAAACTTC 1656
Db 781 GATAAAATTAGAGTAAAAATTTGCTGACCTGGGAAATGCTTGTGTTGGTGATATAAACTTC 840
QY 1657 ACGGAGAGATCCAGAGCGTCAGTACCGCTCCATAGAGTTTTATAGAGCGGGGTAC 1716
Db 841 ACGGAGAGATCCAGAGCGTCAGTACCGCTCCATAGAGTTTTATAGAGCGGGGTAC 900
QY 1717 AGCACCCCTCGGACATCTGAGACAGCGCTGATGGCAATTTGAGCTGGCAACGGGAGAT 1776
Db 901 AGCACCCCTCGGACATCTGAGACAGCGCTGATGGCAATTTGAGCTGGCAACGGGAGAT 960
QY 1777 TATTTGTTGAAACCAATTTCTGGGGAAGACTATTTCCAGAGACGAAACACATAGCCAC 1836
Db 961 TATTTGTTGAAACCAATTTCTGGGGAAGACTATTTCCAGAGACGAAACACATAGCCAC 1020
QY 1837 ATCATAGAGCTGTAGCAGTATTTCCAGGCACTTTGCTCTATCTGGAATAATTTCTCG 1896
Db 1021 ATCATAGAGCTGTAGCAGTATTTCCAGGCACTTTGCTCTATCTGGAATAATTTCTCG 1080
QY 1897 GAATTTCTTCAATCCAGAGGAGAACTGCGACACATCAACCAAGCTGAAGCCCTGGAGCCT 1956
Db 1081 GAATTTCTTCAATCCAGAGGAGAACTGCGACACATCAACCAAGCTGAAGCCCTGGAGCCT 1140
```

```
QY 1957 TTTGATGTAATTTGGAAAAAGTATGCTGGCCCCCATGAAGATGCTGCACAGTTTACAGAT 2016
Db 1141 TTTGATGTAATTTGGAAAAAGTATGCTGGCCCCCATGAAGATGCTGCACAGTTTACAGAT 1200
QY 2017 TTCTGATCCCGATGTAGAAATGGTTCCAGAAAAACGAGCCTCAGCTGGCGAATGCCTT 2076
Db 1201 TTCTGATCCCGATGTAGAAATGGTTCCAGAAAAACGAGCCTCAGCTGGCGAATGCCTT 1260
QY 2077 CGGCATCCTTTGGTTGAATTCCT 2097
Db 1261 CGGCATCCTTTGGTTGAATTCCT 1281

RESULT 2
BX463511 Homo sapiens ADULT BRAIN Homo sapiens linear EST 05-MAY-2004
LOCUS BX463511.2 GI:47056867
DEFINITION CS0DN003YA14 5-PRIME, mRNA sequence.
ACCESSION BX463511
VERSION BX463511.2
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 995)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 22, 2003 this sequence version replaced gi:31027556.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to this cluster, see
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DN003BA07QP1&c=4027.r.

FEATURES
Location/Qualifiers
source
1..995
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN003YA14"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT_6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN
Query Match 46.5%; Score 975.6; DB 5; Length 995;
Best Local Similarity 99.4%; Pred. No. 7.6e-236;
Matches 986; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 820 AAAATATCTTAAACAAAAGAAAACCTGAAAAGAAAACAGAGAGCGCTGAGTTA 879
Db 5 AAAATATCTTAAACAAAAGAAAACCTGAAAAGAAAACAGAGAGCGCTGAGTTA 64
QY 880 TTGGAGAGCGCTCGAGAGATAGAGAAATTTGGAGCGAGAGCTGAAAGGAAAATA 939
Db 65 TTGGAGAGCGCTCGAGAGATAGAGAAATTTGGAGCGAGAGCTGAAAGGAAAATA 124
QY 940 GAAGAAAACATCACCTCAGCTGCACTTCCAAATGACAGGATGGCGAATCTGCCAGAG 999
Db 125 GAAGAAAACATCACCTCAGCTGCACTTCCAAATGACAGGATGGCGAATCTGCCAGAG 184
```

```
QY 1000 CTGAACCTAAACACAGATTAGAGAGGGGGCTGAGCAGAGACTGCAAAAGCAAT 1059
Db 185 GTGAACCTAAACACAGATTAGAGAGGGGGCTGAGCAGAGACTGCAAAAGCAAT 244
QY 1060 GGTGAAGCTGAGCAGCAGAAAGAAAGATGCTGAGAAAGAAACATTTGAAAAGAT 1119
Db 245 GGTGAAGCTGAGCAGCAGAAAGAAAGATGCTGAGAAAGAAACATTTGAAAAGAT 304
QY 1120 GAAGATGATGTAGATCAGAACTTGGCAACATAGACCCCTACGTGGATAGAAATCACTAAA 1179
Db 305 GAAGATGATGTAGATCAGAACTTGGCAACATAGACCCCTACGTGGATAGAAATCACTAAA 364
QY 1180 ACCAATGGCCATATTGAGAAATGGCCCATTTCTACTGGAGCAGCAACTGACCATGAGAT 1239
Db 365 ACCAATGGCCATATTGAGAAATGGCCCATTTCTACTGGAGCAGCAACTGACCATGAGAT 424
QY 1240 GATGATGAAGAGACTGCCCAATCTCTGAGGAATATAATCTTGATGAGCAAAATGCAGAA 1299
Db 425 GATGATGAAGAGACTGCCCAATCTCTGAGGAATATAATCTTGATGAGCAAAATGCAGAA 484
QY 1300 AGTGATTAACATATAGCAGCTCTTATGAACAATTCATGGTGAATTCGCAATGACGCA 1359
Db 485 AGTGATTAACATATAGCAGCTCTTATGAACAATTCATGGTGAATTCGCAATGACGCA 544
QY 1360 CATAAATTCGAGTCAGATCCAGATTCAGAGATTTCCACCTCGTTGTTCTCTGATCCTTA 1419
Db 545 CATAAATTCGAGTCAGATTCAGAGATTTCCACCTCGTTGTTCTCTGATCCTTA 604
QY 1420 GAACCTGTGCGCTGCGGCTCTGTGCTTCTGAGGATACCACTTACTGAGCAAGAGAG 1479
Db 605 GAACCTGTGCGCTGCGGCTCTGTGCTTCTGAGGATACCACTTACTGAGCAAGAGAG 664
QY 1480 AGCAGTCCATCCATGACAGAGCAGACGGTTTCAGCCTCCAGTACTCGGGATTTCGCA 1539
Db 665 AGCAGTCCATCCATGACAGAGCAGACGGTTTCAGCCTCCAGTACTCGGGATTTCGCA 724
QY 1540 AAAGCAAAACCGGCGCAGCTGACTTGTGGTGAATCCCTCGATCCCGCGGAATGCAGAT 1599
Db 725 AAAGCAAAACCGGCGCAGCTGACTTGTGGTGAATCCCTCGATCCCGCGGAATGCAGAT 784
QY 1600 AAAATTAGAGTAAATTCCTGACCTGGGAATGCTTGTGGTGCATAAACCTTCACG 1659
Db 785 AAAATTAGAGTAAATTCCTGACCTGGGAATGCTTGTGGTGCATAAACCTTCACG 844
QY 1660 GAAGACATCCAGACCGCTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGC 1719
Db 845 GAAGACATCCAGACCGCTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGC 904
QY 1720 ACCCTGCGGACATCTGAGGACGCGGCTGTATGGCAATTTGAGCTGCGCAACGGGAGATTAT 1779
Db 905 ACCCTGCGGACATCTGAGGACGCGGCTGTATGGCAATTTGAGCTGCGCAACGGGAGATTAT 964
QY 1780 TTGTTTGACCACTTCTCGGGAAGACTATTC 1811
Db 965 TTGTTTGAA-CMCAATTCGGGGAGACTATTC 995
```

```
RESULT 3
LOCUS BX384358
DEFINITION BX384358 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
cDNA clone CS0DK009YC11 5-PRIME, mRNA sequence.
ACCESSION BX384358
VERSION BX384358.2 GI:46574523
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1001)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
```

```
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30451233.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4027.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DK009AB06QPlc=4027.r.
FEATURES
Location/Qualifiers
1..1001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK009YC11"
/cell_line="HELA"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 44.2%; Score 927.6; DB 5; Length 1001;
Best Local Similarity 97.8%; Pred. No. 1.1e-223;
Matches 960; Conservative 10; Mismatches 9; Indels 3; Gaps 3;
QY 523 ATGGCTCTTCGAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAAACTCAACTATCAA 582
Db 1 ATGGCTCTTCGAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAAACTCAACTATCAA 60
QY 583 GGCCTCCAGTACGTTGTGGAAGATATCATTCGACAGGTCCTTCAAGGTTTAGATTAC 642
Db 61 GGCCTCCAGTACGTTGTGGAAGATATCATTCGACAGGTCCTTCAAGGTTTAGATTAC 120
QY 643 TTACACAGTAAGTGAAGATCATTCATCTGCATATAAGCCGAAATATCTTGATGTGT 702
Db 121 TTACACAGTAAGTGAAGATCATTCATCTGCATATAAGCCGAAATATCTTGATGTGT 180
QY 703 GTGGATGATGCATATGTGAGAAAGATGCGAGCTGAGGCCCACTGAGTGGCAGAAAGCAGGT 762
Db 181 GTGGATGATGCATATGTGAGAAAGATGCGAGCTGAGGCCCACTGAGTGGCAGAAAGCAGGT 240
QY 763 GTCCTCTCTCTTACGGGCTCGAGTACGGTCCACAGCAGAAACCTATAGGAAA 822
Db 241 GTCCTCTCTCTTACGGGCTCGAGTACGGTCCACAGCAGAAACCTATAGGAAA 300
QY 823 ATATCTTAAACCAAAAGAAAGAACTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTTC 882
Db 301 ATATCTTAAACCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTTC 360
QY 883 GAGAAAGCGCTGCAGGAGATAGAAATTTGGAGCCAGAGAGCTGAAAGGAAAATAATAGAA 942
Db 361 GAGAAAGCGCTGCAGGAGATAGAAATTTGGAGCCAGAGAGCTGAAAGGAAAATAATAGAA 420
QY 943 GAAACATCACTCAGTGCACCTTCCAAATGACAGGATGGGGAATCTGCCAGAGGTG 1002
Db 421 GAAACATCACTCAGTGCACCTTCCAAATGACAGGATGGGGAATCTGCCAGAGGTG 480
QY 1003 AAATAAACAACAAGGATTTAGAGGCGGCTGAGGAGAGACTGCAAGGAGCAATGCT 1062
Db 481 AAATAAACAACAAGGATTTAGAGGCGGCTGAGGAGAGACTGCAAGGAGCAATGCT 540
QY 1063 GAAGCTGAGGACGAGGAGAAAGAAAGATGCTGAGAAAGAAAGAAAGAAAGATGAA 1122
Db 541 GAAGCTGAGGACGAGGAGAAAGAAAGATGCTGAGAAAGAAAGAAAGATGAA 600
```

```
JOURNAL Unpublished (2001)
COMMENT On May 8, 2003 this sequence version replaced gi:30451233.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
4027.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DK009AB06QPlc=4027.r.
FEATURES
Location/Qualifiers
1..1001
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK009YC11"
/cell_line="HELA"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 44.2%; Score 927.6; DB 5; Length 1001;
Best Local Similarity 97.8%; Pred. No. 1.1e-223;
Matches 960; Conservative 10; Mismatches 9; Indels 3; Gaps 3;
QY 523 ATGGCTCTTCGAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAAACTCAACTATCAA 582
Db 1 ATGGCTCTTCGAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAAACTCAACTATCAA 60
QY 583 GGCCTCCAGTACGTTGTGGAAGATATCATTCGACAGGTCCTTCAAGGTTTAGATTAC 642
Db 61 GGCCTCCAGTACGTTGTGGAAGATATCATTCGACAGGTCCTTCAAGGTTTAGATTAC 120
QY 643 TTACACAGTAAGTGAAGATCATTCATCTGCATATAAGCCGAAATATCTTGATGTGT 702
Db 121 TTACACAGTAAGTGAAGATCATTCATCTGCATATAAGCCGAAATATCTTGATGTGT 180
QY 703 GTGGATGATGCATATGTGAGAAAGATGCGAGCTGAGGCCCACTGAGTGGCAGAAAGCAGGT 762
Db 181 GTGGATGATGCATATGTGAGAAAGATGCGAGCTGAGGCCCACTGAGTGGCAGAAAGCAGGT 240
QY 763 GTCCTCTCTCTTACGGGCTCGAGTACGGTCCACAGCAGAAACCTATAGGAAA 822
Db 241 GTCCTCTCTCTTACGGGCTCGAGTACGGTCCACAGCAGAAACCTATAGGAAA 300
QY 823 ATATCTTAAACCAAAAGAAAGAACTGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTTC 882
Db 301 ATATCTTAAACCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGTTC 360
QY 883 GAGAAAGCGCTGCAGGAGATAGAAATTTGGAGCCAGAGAGCTGAAAGGAAAATAATAGAA 942
Db 361 GAGAAAGCGCTGCAGGAGATAGAAATTTGGAGCCAGAGAGCTGAAAGGAAAATAATAGAA 420
QY 943 GAAACATCACTCAGTGCACCTTCCAAATGACAGGATGGGGAATCTGCCAGAGGTG 1002
Db 421 GAAACATCACTCAGTGCACCTTCCAAATGACAGGATGGGGAATCTGCCAGAGGTG 480
QY 1003 AAATAAACAACAAGGATTTAGAGGCGGCTGAGGAGAGACTGCAAGGAGCAATGCT 1062
Db 481 AAATAAACAACAAGGATTTAGAGGCGGCTGAGGAGAGACTGCAAGGAGCAATGCT 540
QY 1063 GAAGCTGAGGACGAGGAGAAAGAAAGATGCTGAGAAAGAAAGAAAGAAAGATGAA 1122
Db 541 GAAGCTGAGGACGAGGAGAAAGAAAGATGCTGAGAAAGAAAGAAAGATGAA 600
```

Qy	1123	GATGATGTAGATCAGGAACCTGGCAACATAGACCTTACGTGGATAGAAATCACCTTAAACC	1182
Db	601	GATGATGTAGATCAGGAACCTGGCAACATAGACCTTACGTGGATAGAAATCACCTTAAACC	660
Qy	1183	AATGGCCATATTGAGAAATGGCCCAATTCCTACTGGAGCAGCAACTGGACGATGAAGATGAT	1242
Db	661	AATGGCCATATTGAGAAATGGCCCAATTCCTACTGGAGCAGCAACTGGACGATGAAGATGAT	720
Qy	1243	GATGAAGAAGACTGCCCAATTCCTTGAGGAATATAATCTTGATGAGCCAAATGCAGAAAAGT	1302
Db	721	GATGAAGAAGAAATGCCCAATTCCTTGAGGAATATAATCTTGATGAGCCAAATGCAGAAAAGT	780
Qy	1303	GATTACACATATAGCAGCTCTATGAAACAATCAATGGTGAATTCCTCAAAATGGACGACAT	1362
Db	781	GATTACACATATAGCAGCTCTATGAAAMAATTCAAATGGTGAATTCCTCAAAATGGACGACAT	840
Qy	1363	AAAATTCGGAGTCACAGTTTCCAGAGTTTTCCACCTCTGTTGTTCTCTGGATCCTTAGAA	1422
Db	841	AAAATTCGGAGTCACAGTTTCCAGAGTTTTCCACCTCGTTGTTCTCTGGATCCTTAGAA	900
Qy	1423	CCTGTGGCCTCGGGCTCTGTGCTTTTCTGAGGGATCAACACTTTACTGACGAAGAGGAGAGC	1482
Db	901	CCTGTGGCCTCGGG-TCCTGTGCYTTCTGAGGGAT-MCCMCTTACTGTGRC-AGRGGAGGC	957
Qy	1483	AGTCCATCCCATGACAGAGCA	1504
Db	958	AGTCATCCCGACAGAAAMRCA	979

RESULT	4
BM464185	
LOCUS	BM464185 1108 bp linear mRNA EST 05-FEB-2002
DEFINITION	AGENCOURT_6438899 NIH_MGC_71 Homo sapiens cdna clone IMAGE:5535256 5', mRNA sequence.

EST. Homo sapiens (human)
KEYWORDS Homo sapiens
SOURCE
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1108)
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLAM13223 row: a column: 17
high quality sequence stop: 787.

```

FEATURES
  source
    high quality sequence
    Location/Qualifiers
      1..1108
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5535256"
        /tissue_type="leiomyosarcoma"
        /lab_host="DH10B (phage-resistant)"
        /clone_lib="NIH MGC 71"
        /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
        Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
        Average insert size 2.1 kb."
ORIGIN
  Query Match      38.9%;   Score 816.2;   DB 4;   Length 1108;
  Best Local Similarity 99.6%;   Pred. No. 2e-195;
  Matched e.g. Conservative 0;   Mismatches 3;   Indels 0;   Gaps 0;

```

	QY	1277	ATCTTGATGAGCCAAATGCAGAAAGTGAATTACACATATAGCAGCTCCCTATGAAACAATTC	1336
	Dd	21	ATCTTGATGAGCCAAATGCAGAAAGTGAATTACACATATAGCAGCTCCCTATGAAACAATTC	80
	QY	1337	ATCGTGAATTTGCCAATGGACACATATAAATTTCCGAGTCAACAGTTCCGAGAGTTTCCA	1396
	Dd	81	ATGGTGAATTTGCCAATGGACACATATAAATTTCCGAGTCAACAGTTCCGAGAGTTTCCA	140
	QY	1397	CCTCGTTGTCTCTGGAATCCTTAGAACCTGTGGCTGCGGCTCTGTGCTTTCTGAGGGAT	1456
	Dd	141	CCTCGTTGTCTCTGGAATCCTTAGAACCTGTGGCTGCGGCTCTGTGCTTTCTGAGGGAT	200
	QY	1457	CACCATTACTGAGCAAGGAGCAGTCCATCCCATGACAGAAGCAGAACGGTTTTCTAG	1516
	Dd	201	CACCATTACTGAGCAAGGAGCAGTCCATCCCATGACAGAAGCAGAACGGTTTTCTAG	260
	QY	1517	CCTCCAGTACTCGGGATTTCCAAAAAGCAAAAACCCGGCAGCTGACTTGTGGTGAATC	1576
	Dd	261	CCTCCAGTACTCGGGATTTCCAAAAAGCAAAAACCCGGCAGCTGACTTGTGGTGAATC	320
	QY	1577	CCCTGGATCCCGGAATGCAAGATAAAATTTAGAGTAATAATTTCTGACCTCGGGAATGCTT	1636
	Dd	321	CCCTGGATCCCGGAATGCAAGATAAAATTTAGAGTAATAATTTCTGACCTCGGGAATGCTT	380
	QY	1637	GTTGGGTGCATAAACACTTCCACGGAGACATCCAGACGGCTCAGTACCGCTCCATAGAG	1696
	Dd	381	GTTGGGTGCATAAACACTTCCACGGAGACATCCAGACGGCTCAGTACCGCTCCATAGAG	440
	QY	1697	TTTTAATAGGACGGGGTACAGCACCCCTCGGACATCTGGAGCAGCGGTGATGGCAT	1756
	Dd	441	TTTTAATAGGACGGGGTACAGCACCCCTCGGACATCTGGAGCAGCGGTGATGGCAT	500
	QY	1757	TGAGCTGGGAAACGGGAGATTATTTGTTTGAACACACTTCTGGGGAAGACTATTTCCAGAG	1816
	Dd	501	TGAGCTGGGAAACGGGAGATTATTTGTTTGAACACACTTCTGGGGAAGACTATTTCCAGAG	560
	QY	1817	ACGAAGACCATATAGCCACATCATAGAGTGTCTAGGACGATTTTCCAAGCAGCTTTGCTC	1876
	Dd	561	ACGAAGACCATATAGCCACATCATAGAGTGTCTAGGACGATTTTCCAAGCAGCTTTGCTC	620
	QY	1877	TATCTGGAAAAATTTCTCGGAAATCTTCAAATCCGACAGAGAACTCGGACACATACCA	1936
	Dd	621	TATCTGGAAAAATTTCTCGGAAATCTTCAAATCCGACAGAGAACTCGGACACATACCA	680
	QY	1937	AGCTGAAGCCCTGGAGCCTCTTTCATGTPACTTGTGGAAAAAGTATGGCTGGCCCCATGAAG	1996
	Dd	681	AGCTGAAGCCCTGGAGCCTCTTTCATGTPACTTGTGGAAAAAGTATGGCTGGCCCCATGAAG	740
	QY	1997	ATGCTGCACAGTTTACAGATTTCTGTATCCCAGATTTTAGAAAAATGGTTCCAGAAAAACGAG	2056
	Dd	741	ATGCTGCACAGTTTACAGATTTCTGTATCCCAGATTTTAGAAAAATGGTTCCAGAAAAACGAG	800
	QY	2057	CCTCAGCTGGCGAATGCCCTTCGGCATCCTTGGTTGAATTC	2097
	Dd	801	CCTCAGCTGGCGAATGCCCTTCGGCATCCTTGGTTGAATTC	841

RESULT 5
 BU166874
 LOCUS
 DEFINITION

BU166874
 AGENCOURT_7960371 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:616382
 5', mRNA sequence.

linear EST 04-SEP-2002

RESULT 5	ACCESSION	REFERENCE
BU166874	VERSION	AUTHOR
LOCUS	KEYWORDS	TITLE
DEFINITION	SOURCE	
	ORGANIZATION	

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTP/DTP
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM13526 Row: j Column: 15
 High quality sequence stop: 579.
 Location/Qualifiers
 1. .867
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6166382"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 72"
 /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally, primer: Oligo dt.
 Average insert size 2 kb. Library constructed by Life
 Technologies."

ORIGIN

Query Match	37.4%	Score 785;	DB 5;	Length 867;
Best Local Similarity	96.1%	Pred. No. 1.5e-187;		
Matches 827;	Conservative 0;	Mismatches 30;	Indels 4;	Gaps 2;
Qy	1120	GAAGATGATGTAGATCAGGAACTTGC	GAACATAGACACCTCAGTGGATAGATCACTAAA	1179
Db	1	GAAGATGATGTAGATCAGGAACTTGC	GAACATAGACACCTCAGTGGATAGATCACTAAA	60
Qy	1180	ACCAATGGCCATATTGAGATGGCCCATTTCT	CATCTGGAGCAGCAACTGGACGATGAAGAT	1239
Db	61	ACCAATGGCCATATTGAGATGGCCCATTTCT	CATCTGGAGCAGCAACTGGACGATGAAGAT	120
Qy	1240	GATCATGAAGAAGACTGCCAAATCTCTGAGGAATATATCT	TGTATGAGCCAAATGCAGAA	1299
Db	121	GATCATGAAGAAGACTGCCAAATCTCTGAGGAATATATCT	TGTATGAGCCAAATGCAGAA	180
Qy	1300	AGTGATTACACATATAGCAGCTCCTCATGAACAATTC	AATTCGATGGTGAATTCGCCAAATGGACGA	1359
Db	181	AGTGATTACACATATAGCAGCTCCTCATGAACAATTC	AATTCGATGGTGAATTCGCCAAATGGACGA	240
Qy	1360	CATAAATTTCCGAGTACAGTTC	CCAGAGTTCCTCAGCTCGTTGTTCTCTGGATTCCTTA	1419
Db	241	CATAAATTTCCGAGTACAGTTC	CCAGAGTTCCTCAGCTCGTTGTTCTCTGGATTCCTTA	300
Qy	1420	GAACCTGTGGCTCGGGCTGTGCTTTCTGAGGGATCACAC	TACTGTGACGAGGAG	1479
Db	301	GAACCTGTGGCTCGGGCTGTGCTTTCTGAGGGATCACAC	TACTGTGACGAGGAG	360
Qy	1480	AGCAGTCCATCCCATGACAGAAGCAGAACGGTTT	CAGCCTCCAGTCTGGGGATTTGCCA	1539
Db	361	AGCAGTCCATCCCATGACAGAAGCAGAACGGTTT	CAGCCTCCAGTCTGGGGATTTGCCA	420
Qy	1540	AAAGCAAAACCCGGGAGCTGACTTGTGTGTGAATCCCT	TGGATCCGCGGAATGCAGAT	1599
Db	421	AAAGCAAAACCCGGGAGCTGACTTGTGTGTGAATCCCT	TGGATCCGCGGAATGCAGAT	480
Qy	1600	AAATATTAGCTAAAATTTGCTGACCTGGGAATGCTTGT	TGGTGGTGCATTAACACTTCCAG	1659
Db	481	AAATATTAGCTAAAATTTGCTGACCTGGGAATGCTTGT	TGGTGGTGCATTAACACTTCCAG	540
Qy	1660	GAAGACATCCAGACCGCTCAGTACCGCTCCATAGAGGTTT	TAAATAGGAGCGGGTACAGC	1719
Db	541	GAAGACATCCAGACCGCTCAGTACCGCTCCATAGAGGTTT	TAAATAGGAGCGGGTACAGC	600
Qy	1720	ACCCTTCGGACATCTGGAGCACCGGCTGTATTGCA	TTTGAGCTGGCAACGGGAGATTAT	1779

1414 TCCTTAGAACCTGTGGCTGGCGCTCTGTGCTTTCTGAGGATCACCACCTTACTGAGCAA 1473
 |||||
 661 TCCTTAGAACCTGTGGCTGGCGCTCTGTGCTTTCTGAGGATCACCACCTTACTGAGCAA 720
 |||||
 1474 GA-GGAGACAGTCCATCCCATGACAGAAGCAGAAC-GGTTTCAGCTCCAG--TACTGG 1529
 |||||
 721 GAGGAGAGACAGTCCATCCCTGACAGAAGCCGACCGGTTTCAGCTCCAGGACTGGG 780
 |||||
 1530 GGATTTGCCAAA--GCAAAAACCGGGCAGCTGACTTGT--GGTGAATCCCTGGATC 1585
 |||||
 781 GGATTTGCCAAAAGCAAAAACCCCGGCAGCTGACCGGTTTGGTGAATCCCTGGATC 840
 |||||
 1586 CGC-GGAATGCAGATAAAA--TTAGACTAAAAATTG-CTGACCTGGGAAA---TGCTTGT 1639
 |||||
 841 CCGGGAGCGAGATAAAAATTAGATAAAAATTGCTGACCTGAGAAAATGCTTGGTG 900
 |||||
 1640 GGGTGCAAAACACTTTCACGGAA 1662
 |||||
 901 GGGGGCATAAACCTTTCCGGAA 923
 |||||

RESULT 8
 BX423898
 LOCUS
 DEFINITION BX423898 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
 CS0DN003YA14 5-PRIME, mRNA sequence.
 ACCESSION BX423898
 VERSION BX423898.2 GI:46927093
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30643492.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4027.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?b=CSIAN001ZE08QPl&c=4027.r.

FEATURES
 source
 1. .900
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DN003YA14"
 /tissue_type="ADULT BRAIN"
 /dev_stage="adult"
 /clone_lib="Homo sapiens ADULT BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN
 Query Match 35.4%; Score 741.8; DB 5; Length 900;
 Best Local Similarity 97.0%; Pred. No. 1.4e-176;
 Matches 808; Conservative 2; Mismatches 14; Indels 9; Gaps 5;
 820 AAAATATCTAAAAACAAAAGAAAAAAGTGAAGAAAAAGAAAGAGGAGGCTGAGTTA 879
 |||||
 5 AAATATCTAAAAACAAAAGAAAAAAGTGAAGAAAAAGAAAGAGGAGGCTGAGTTA 64
 |||||

880 TTGGAGAGC--GCCTCAGGAGATAGAGAAATTCGAGCGAGAGCTGAAAGAAATAA 937
 |||||
 65 TTGGAGAGC--GCCTCAGGAGATAGAGAAATTCGAGCGAGAGCTGAAAGAAATAA 124
 |||||
 938 TAGAAGAAAAACATCACTCAGCTGCACTTCCAAATGACCAGGATGCGGAATCTGCCAG 997
 |||||
 125 TAGAAGAAAAACATCACTCAGCTGCACTTCCAAATGACCAGGATGCGGAATCTGCCAG 184
 |||||
 998 AGGTGAACATAAATAAACAACAGATTAGAGAGCGGCTGAGCAGAGATCTGCAAGAGCA 1057
 |||||
 185 AGGTGAACATAAATAAACAACAGATTAGAGAGCGGCTGAGCAGAGATCTGCAAGAGCA 244
 |||||
 1058 ATGCTGACCTGAGCAGCAGGAGAGAGAGATCTCAGAAAGAAACATTTGAAAG 1117
 |||||
 245 ATGGTGAAGCTGAGCAGCAGGAGAGAGATCTCAGAAAGAAACATGAGAAAAAG 304
 |||||
 1118 ATGAAGATGATGTAGATCAGGAACCTTCGAAACATAGACCTACCTGCGATAGATCACCTA 1177
 |||||
 305 ATGAAGATGATGTAGATCAGGAACCTTCGAAACATAGACCTACCTGCGATAGATCACCTA 364
 |||||
 1178 AAACCAATGGCCATATTGAGAAATGCCCATTTCTACTGAGCAGCAACTGGACGATGAAG 1237
 |||||
 365 AAACCAATGGCCATATTGAGAAATGCCCATTTCTACTGAGCAGCAACTGGACGATGAAG 424
 |||||
 1238 ATGATGATGAAGAAGACTGCCCAAAATCCTGAGGAATATTAATCTTGATGAGCCAAATGCG 1297
 |||||
 425 ATGATGATGAAGAAGACTGCCCAAAATCCTGAGGAATATTAATCTTGATGAGCCAAATGCG 484
 |||||
 1298 AAGTGATTACACATATATAGCAG-CTCCTATGAACAATCAATGGTGAATTTGCCAAATGGA 1356
 |||||
 485 AAGTGATTACACATATATAGCAGCTCCTATGAACAATCAATGGTGAATTTGCCAAATGGA 544
 |||||
 1357 CGACATAAAATTCGCGAGTCACAGTTCCAGAGATTTCCACCTCGTTGTTCTCTGGATCC 1416
 |||||
 545 CGACATAAAATTCGCGAGTCACAGTTCCAGAGATTTCCACCTCGTTGTTCTCTGGATCC 604
 |||||
 1417 TTAGAACCTGTGGCTGGCGCTCTGTGCTTTCTGAGGATCACCACCTTACTGAGCAAGAG 1476
 |||||
 605 TTAGAACCTGTGGCTGGCGCTCTGTGCTTTCTGAGGATCACCACCTTACTGAGCAAGAG 664
 |||||
 1477 GAGACAGTCCATCCATGACAGAAGCAGAACGGTTTCAGCTCCAGTACTGGGATTTG 1536
 |||||
 665 GAGACAGTCCATCCATGACAGAAGCAGAACGGTTTCAGCTCCAGTACTGGGATTTG 724
 |||||
 1537 CCAAAAGCAAAACCCGGGCGAGCTGACTTTGGTGAATCCCTCGATCCGCGGAATGCA 1596
 |||||
 725 CCAAAAGCAAAACCCGGGCGAGCTGACTTTGGTGAAT-CCCTGGATCCGCGG-ATGCA 782
 |||||
 1597 GATAAAATTAGAGTAAAAAATTCCTGACCTGGGAAATGCTTGTGGTGGTGAATAA 1649
 |||||
 783 GATAAAATTAGAGTAAAAAATTTTG----ACTKGGGAATGCTTGTGGTGGTGAATAA 831
 |||||

RESULT 9
 BX442846
 LOCUS
 DEFINITION BX442846 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
 CS0DF038YI10 5-PRIME, mRNA sequence.
 ACCESSION BX442846
 VERSION BX442846.2 GI:47037734
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 22, 2003 this sequence version replaced gi:31018541.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France
 Email: seque@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
 end enriched double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4027.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?c=CS0DF038BE05QP1&c=4027.r.

Location/Qualifiers
 1. .960
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF038Y110"
 /dev_stage="fetal"
 /tissue_type="FETAL BRAIN"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo (dT) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

ORIGIN

Query Match 35.2%; Score 739; DB 5; Length 960;
 Best Local Similarity 94.5%; Pred. No. 7,3e-176;
 Matches 757; Conservative 25; Mismatches 16; Indels 3; Gaps 2;

QY 63 TCCGAAAGCCGGAGCCTCAACAGAAAGCTCCTTAGTTCCTCTCTCCACCCACC 122
 Db 154 TTCAAGAGCCGGAGCCTCAACAGAAAGCTCCTTAGTTCCTCTCTCCACCCACC 213

QY 123 ACCACACC-ACCGCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGATC 180
 Db 214 ACCACACCAGCGCGCCACTTTGCCAGACCCACACCCCGGAGCCAGAGGAGATC 273

QY 181 CTGGGATCAGATGATGAGGAGCAAGAGGACCTTGGGAGTACTGCAAAAGGTGGATATCAT 240
 Db 274 CTGGGATCAGATGATGAGGAGCAAGAGGACCTTGGGAGTACTGCAAAAGGTGGATATCAT 333

QY 241 CCAAGTAAATTTGGAGACCTTTCAATGCGCGGTATCATGTTATTAGAAAGCTTGGATGG 300
 Db 334 CCAAGTAAATTTGGAGACCTTTCAATGCGCGGTATCATGTTATTAGAAAGCTTGGATGG 393

QY 301 GGGCACTTCTACTGTCTGCTGGTGTCTGGGATGTCGAGGCAAGAGTTTCTTCAATG 360
 Db 394 GGGCACTTCTACTGTCTGCTGGTGTCTGGGATGTCGAGGCAAGAGTTTCTTCAATG 453

QY 361 AAAGTTGTAAGAGTCCCGAGCATTTATACGGAGACAGCCCTTGGATGAAATAAAATTTGCTC 420
 Db 454 AAAGTTGTAAGAGTCCCGAGCATTTATACGGAGACAGCCCTTGGATGAAATAAAATTTGCTC 513

QY 421 AAATGTTTCGAGAAAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCC 480
 Db 514 AAATGTTTCGAGAAAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCCAGTATCC 573

QY 481 GAGCACTTCAGATTTTCAGGATGATGAGTATGATGATGATGATGATGATGATGATGATGATGAT 540
 Db 574 GAGCACTTCAGATTTTCAGGATGATGAGTATGATGATGATGATGATGATGATGATGATGATGAT 633

QY 541 GGCACCATCTCTCTCAAGTGGATCATCAATCAACTATCAAGGCTCCAGTACGTTGT 600
 Db 634 GGCACCATCTCTCTCAAGTGGATCATCAATCAACTATCAAGGCTCCAGTACGTTGT 693

QY 601 GTGAAGAGTATCATTTTCAGAGGTCCTTCAAGGTTAGATTAATTACAGAGTAAAGTGAAG 660
 Db 694 GTGAAGAGTATCATTTTCAGAGGTCCTTCAAGGTTAGATTAATTACAGAGTAAAGTGAAG 753

QY 661 ATCATTTTCATCTGACATTAAGCCGAAATATCTTGTGATGATGATGATGATGATGATGATGATGAT 720
 Db 754 ATCATTTTCATCTGACATTAAGCCGAAATATCTTGTGATGATGATGATGATGATGATGATGATGAT 813

QY 721 AGAAGATGCGAGCTGAGGCGCAGTGTGCGAGAAAGCAGGTGCTCTCTCTCTTCA-GG 779
 Db 814 AGAAGATGCGAGCTGAGGCGCAGTGTGCGAGAAAGCAGGTGCTCTCTCTCTTCA-RRG 873

QY 780 GTCTGCGAGTGTGAGTGTGCGCTCCAGCAGAGAAACCTATAGGAAATATCTAAACAAAAA 839
 Db 874 GTTTCAGTGTGAGTGTGCGCTCCAGCAGAGAAACCTATAGGAAATATCTAAACAAAAA 933

QY 840 GAAAAAACTGAAAAAGAAACA 860
 Db 934 RAAAAACYGAAAAAACAARA 954

RESULT 10
 BU607396/c
 LOCUS
 DEFINITION
 UI-CF-FNO-aeu-b-08-0-UI.s1 UI-CF-FNO Homo sapiens cDNA clone
 UI-CF-FNO-aeu-b-08-0-UI 3', mRNA sequence.

ACCESSION
 VERSION
 BU607396
 BU607396.1 GI:23270671
 EST.
 KEYWORDS
 SOURCE
 Homo sapiens (human)

ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 735)

AUTHORS
 Bonaldo,M.F., Lennon,G. and Soares,M.B.

TITLE
 Normalization and subtraction: two approaches to facilitate gene

JOURNAL
 MEDLINE
 Genome Res. 6 (9), 791-806 (1996)

PUBMED
 97044477

COMMENT
 8889548
 Contact: McCray, PB

McCray Lab
 University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

Cloning by: Dr. M. Bento Soares, University of Iowa

Cloning Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com) or from Open Biosystems

(www.openbiosystems.com).

The following repetitive elements were found in this cDNA

sequence: 686-735, >(TGG)n#Simple_repeat

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

1. .735

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-CF-FNO-aeu-b-08-0-UI"

/tissue_type="Human lung Epithelial cells"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_libs="UI-CF-FNO"

/note="Organ: Lung; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site_1: EcoR I; Site_2: Not I;

UI-CF-FNO is a subtracted cDNA library derived from two

normalized Human lung epithelial cell libraries (EN1 and

DUI) The library was subtracted according to according to

Bonaldo, Lennon and Soares, Genome Research, 6:791-806,

1996. For additional information, contact:

bento-soares@uiowa.edu

TAG TISSUE=Human Lung Epithelial Cell Lines untreated LPS

6hr to LPS 24h

TAG LIB=UI-CF-FNO

TAG_SEQ=CTGCTCAGGT"

ORIGIN

Query Match	34.8%	Score 730.4	DB 5	Length 735	
Best Local Similarity	99.5%	Pred. No. Id=173			
Matches 731	Conservative 0	Mismatches 4	Indels 0	Gaps 0	
112	CCACCGCCACACACACACACCGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAG 171				
735	CCACCGCCACACACACACCGCCACCTNTTGCAGACCCACACCCCGGAGCCAGAG 676				
172	GAGGAGATCTGGGATCAGATGATGAGGAGCAGAGGACCTCTGGGATCTATGCAAGGT 231				
675	GAGGAGATCTGGGATCAGATGATGAGGAGCAGAGGACCTCTGGGATCTATGCAAGGT 616				
232	GGATATCATCCAGTCAAAATTTGGAGACCTCTTCAATGCCGATATCATGTTATTAGAAG 291				
615	GGATATCATCCAGTCAAAATTTGGAGACCTCTTCAATGCCGATATCATGTTATTAGAAG 556				
292	CTTGATGGGGCATCTCTTACTGTCTGGCTGTCTGGGATATGACGGGGAAAAAGTTT 351				
555	CTTGATGGGGCATCTCTTACTGTCTGGCTGTCTGGGATATGACGGGGAAAAAGTTT 496				
352	GTTCGAATGAAAGTTGTAAGAAGTCCCGACGATATATCGGAGACAGCCTTCGGATGAATA 411				
495	GTTCGAATGAAAGTTGTAAGAAGTCCCGACGATATATCGGAGACAGCCTTCGGATGAATA 436				
412	AAATTGCTCAAAATGTTTCGAGAAAGTGATCCAGTGACCCCAAAACAAAGACATGGTGTC 471				
435	AAATTGCTCAAAATGTTTCGAGAAAGTGATCCAGTGACCCCAAAACAAAGACATGGTGTC 376				
472	CAGCTCATGTGACGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTCGATGGTCTTC 531				
375	CAGCTCATGTGACGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTCGATGGTCTTC 316				
532	GAAATGACTTGGCCACCATCTCTCTCAAGTGGATCATCAAAATCCAATCATCAAGGCCCTCCA 591				
315	GAAATGACTTGGCCACCATCTCTCTCAAGTGGATCATCAAAATCCAATCATCAAGGCCCTCCA 256				
592	GTAAGTTGTGGAAGAGTATCATTCGACAGGTCTCTTCAAGGGTTAGATTCTTACACAGT 651				
255	GTACGTTGTGGAAGAGTATCATTCGACAGGTCTCTTCAAGGGTTAGATTCTTACACAGT 196				
652	AAGTCAAGATCATTCATCATCTGACATAAAGCCGAAAAATATCTTGTGCTGTGTGATGAT 711				
195	AAGTCAAGATCATTCATCATCTGACATAAAGCCGAAAAATATCTTGTGCTGTGTGATGAT 136				
712	GCATATGTGAGAAGATGGCAGCTGAGGCCACTGATGGCAGAAAGACAGGTGCTCTCCT 771				
135	GCATATGTGAGAAGATGGCAGCTGAGGCCACTGATGGCAGAAAGACAGGTGCTCTCCT 76				
772	CCTTCAGGCTCTGCAGTACGCTCCACAGCAGAAACCTATAGGAAAAATATCTTAAA 831				
75	CCTTCAGGCTCTGCAGTACGCTCCACAGCAGAAACCTATAGGAAAAATATCTTAAA 16				
832	AACAAAAAGAAAAA 846				
15	AACAAAAAGAAAAA 1				

RESULT 11	
C2242242860/c	
CCUCS	
DEFINITION	728 bp mRNA linear EST 12-FEB-2003
UI-CF-PN0-agb-e-16-0-UI.s1	UI-CF-PN0 Homo sapiens cDNA clone
UI-CF-PN0-agb-e-16-0-UI 3'	mRNA sequence.
CB242860	
CB242860	
CB242860.1	GI:28364504
EST.	
SOURCE	Homo sapiens (human)
TISSUE	Homo sapiens
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 728)
AUTHORS	Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE	Normalization and subtraction: two approaches to facilitate gene

discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
Seq primer: M13 FORWARD

```

FEATURES      Location/Qualifiers
1..728
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="UI-CF-FNO-agb-e-16-0-UI"
    /tissue_type="Human Lung Epithelial cells"
    /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
    /clone_lib="UI-CF-FNO"
    /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-FNO is a subtracted cDNA library derived from
normalized Human lung epithelial cell libraries (EN1 and
DUI) The library was subtracted according to according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. For additional information, contact:
bento-soares@uiowa.edu
TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG LIB=UI-CF-FNO
TAG SEO=GGCTGTAGGC"

```

ORIGIN	Query Match	34.5%;	Score 723.2;	DB 6;	Length 728;
	Best Local Similarity	99.6%;	Pred. No. 6,7e-172;		
	Matches 725;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
Qy	119	CACACACACACACCGCCACCTTCCGAGACCCACACCCCGGAGCCAGAGGAGAGA	178		
Db	728	CACACACACACACCGCCACCTTGTGCAGACCCACACCCCGGAGCCAGAGGAGAGA	669		
Qy	179	TCTTGGGATCAGATGATAGGAGCAAGAGAGCCCTCGGGACTACTCCAAAGGTGGATATC	238		
Db	668	TCTTGGGATCAGATGATAGGAGCAAGAGAGCCCTCGGGACTACTCCAAAGGTGGATATC	609		
Qy	239	ATCCAGTGAANAATTGGAGACCTCTTCAATGGCCGGTATCATGTATTAGAAAGCTTGGAT	298		
Db	608	ATCCAGTGAANAATTGGAGACCTCTTCAATGGCCGGTATCATGTATTAGAAAGCTTGGAT	549		
Qy	299	GGGGGCACCTTCTCTACTGTCTGGCTGTCTGGGATATGACGGGGAAAGATTTGTTCGAA	358		
Db	548	GGGGGCACCTTCTCTACTGTCTGGCTGTCTGGGATATGACGGGGAAAGATTTGTTCGAA	489		
Qy	359	TGAAAGTTGTAAAGAGTGCCACGATTATACGGAGACAGCCTTGGATGAAATTAATTTGC	418		
Db	488	TGAAAGTTGTAAAGAGTGCCACGATTATACGGAGACAGCCTTGGATGAAATTAATTTGC	429		
Qy	419	TCAATATGTGTTTCAGAGAAAGTGATCCACAGTGACCCAAACAAAGACATGCGTCCAGCTCA	478		
Db	428	TCAATATGTGTTTCAGAGAAAGTGATCCACAGTGACCCAAACAAAGACATGCGTCCAGCTCA	369		
Qy	479	TTGACGATTCGAAGATTTTCAGGCATGAATGGGATACATGCTCGATGGTCTTCGAAGTAC	538		
Db	368	TTGACGATTCGAAGATTTTCAGGCATGAATGGGATACATGCTCGATGGTCTTCGAAGTAC	309		


```
QY 539 TTGGCCACCATCTCTCAAGTGGATCATCAATCAACTATCAAGGCTCCAGTACGTT 598
|
|
|
Db 308 TTGGCCACCATCTCTCAAGTGGATCATCAATCAACTATCAAGGCTCCAGTACGTT 249
|
|
|
QY 599 GTGTGAAGAGTATCATTTGACAGAGTCTTCAAGGTTAGATTACTTACACAGTAAGTGA 658
|
|
|
Db 248 GTGTGAAGAGTATCATTTGACAGAGTCTTCAAGGTTAGATTACTTACACAGTAAGTGA 189
|
|
|
QY 659 AGATCATTCATACATGACATTAAGCGGAAATATCTTGTATGTTGTGATGATGATG 718
|
|
|
Db 188 AGATCATTCATACATGACATTAAGCGGAAATATCTTGTATGTTGTGATGATGATG 129
|
|
|
QY 719 TCAGAAGAGTATGACATTAAGCGGAAATATCTTGTATGTTGTGATGATGATG 778
|
|
|
Db 128 TCAGAAGAGTATGACATTAAGCGGAAATATCTTGTATGTTGTGATGATGATG 69
|
|
|
QY 779 GGTCTGCGAGTATGACATTAAGCGGAAATATCTTGTATGTTGTGATGATGATG 838
|
|
|
Db 68 GGTCTGCGAGTATGACATTAAGCGGAAATATCTTGTATGTTGTGATGATGATG 9
|
|
|
QY 839 AGAAAAA 846
|
|
|
Db 8 AAAAAAA 1
```

```
RESULT 12
BM979654/c
LOCUS
DEFINITION
BM979654
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 733)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
97044477
8889548
Contact: McCray, PB
McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com) or from Open Biosystems
(www.openbiosystems.com).
The following repetitive elements were found in this cDNA
sequence: 1-28, >AT-rich#Low_complexity (matched complement)
707-733, >(TGG)n#Simple_repeat
Seq primer: M13 FORWARD
POLVA=Yes.
FEATURES
source
1..733
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-CF-DUI-adt-e-02-0-UI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
```

```
ORIGIN
Query Match 34.2%; Score 717.8; DB 5; Length 733;
Best Local Similarity 99.5%; Pred. No. 1.6e-170;
Matches 730; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
QY 113 CACGCCACACACACACACACCTTTGCCAGACCCACACACCCCGGAGCCAGG 172
|
|
|
Db 733 CACGCCACACACACACACACCTTTGCCAGACCCACACACCCCGGAGCCAGG 675
|
|
|
QY 173 AGGAGATCTCGGATCAGATGATGAGGAGAGGACCTTCCGAGCTACTTGCAGAGTG 232
|
|
|
Db 674 AGGAGATCTCGGATCAGATGATGAGGAGAGGACCTTCCGAGCTACTTGCAGAGTG 615
|
|
|
QY 233 GATATCATCCAGTGAAATTTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGC 292
|
|
|
Db 614 GATATCATCCAGTGAAATTTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGC 555
|
|
|
QY 293 TTGATGGGGACCTCTCTACTCTGGCTGCTGGGATATCAGGGGAAAGATTTG 352
|
|
|
Db 554 TTGATGGGGACCTCTCTACTCTGGCTGCTGGGATATCAGGGGAAAGATTTG 495
|
|
|
QY 353 TTGCAATGAAAGTTGTAAGAGTCCAGCATTTATACGAGACAGCTTGGATGAAATPAA 412
|
|
|
Db 494 TTGCAATGAAAGTTGTAAGAGTCCAGCATTTATACGAGACAGCTTGGATGAAATPAA 435
|
|
|
QY 413 AATTGCTCAATTTGTTGAGAAAGTATCCAGTACCCAAACAAAGACATGTTGGTCC 472
|
|
|
Db 434 AATTGCTCAATTTGTTGAGAAAGTATCCAGTACCCAAACAAAGACATGTTGGTCC 375
|
|
|
QY 473 AGCTCATTTGACGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTTGCATGGTCTTCG 532
|
|
|
Db 374 AGCTCATTTGACGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTTGCATGGTCTTCG 315
|
|
|
QY 533 AAGTACTTGGCCACCATCTCTCAAGTGGATCATCAATCCAACTATCAAGGCTCCAG 592
|
|
|
Db 314 AAGTACTTGGCCACCATCTCTCAAGTGGATCATCAATCCAACTATCAAGGCTCCAG 255
|
|
|
QY 593 TACGTTGTGAAGAGTATCATTTGACAGGTCCTTCAAGGTTAGATTACTTACACAGTA 652
|
|
|
Db 254 TACGTTGTGAAGAGTATCATTTGACAGGTCCTTCAAGGTTAGATTACTTACACAGTA 195
|
|
|
QY 653 AGTCAAGATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 712
|
|
|
Db 194 AGTCAAGATCATTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 135
|
|
|
QY 713 CATATGTGAGAAGATGCGAGCTGAGGCTGAGTGGCAGAAAGACAGGTCCTCTCTC 772
|
|
|
Db 134 CATATGTGAGAAGATGCGAGCTGAGGCTGAGTGGCAGAAAGACAGGTCCTCTCTC 75
|
|
|
QY 773 CTTGAGGTTCTGAGTGTAGTACGGCTCCAGAGCAGAAACCTATAGGAAAAATATCTAAA 832
|
|
|
Db 74 CTTGAGGTTCTGAGTGTAGTACGGCTCCAGAGCAGAAACCTATAGGAAAAATATCTAAA 15
|
|
|
QY 833 ACAAAAAAGAAAAA 846
```

```
/clone_lib="UI-CF-DUI"
/notes="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an EcoR I adaptor, digested with Not I, and cloned
directionally into p7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_LIB=UI-CF-DUI
TAG_SEQ=GGCTGTAGGC"
```

Db	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	TITLE	JOURNAL	COMMENT	FEATURES	source
14	AAAAAAAAAAAAAAAAA												
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
LOCUS	AI038250	711 bp mRNA linear EST 30-JUN-1998											
DEFINITION	AI038250	711 bp mRNA linear EST 30-JUN-1998											
ACCESSION	AI038250	711 bp mRNA linear EST 30-JUN-1998											
VERSION	AI038250	711 bp mRNA linear EST 30-JUN-1998											
KEYWORDS	AI038250	711 bp mRNA linear EST 30-JUN-1998											
SOURCE	AI038250	711 bp mRNA linear EST 30-JUN-1998											
ORGANISM	AI038250	711 bp mRNA linear EST 30-JUN-1998											
REFERENCE	AI038250	711 bp mRNA linear EST 30-JUN-1998											
AUTHORS	AI038250	711 bp mRNA linear EST 30-JUN-1998											
TITLE	AI038250	711 bp mRNA linear EST 30-JUN-1998											
JOURNAL	AI038250	711 bp mRNA linear EST 30-JUN-1998											
COMMENT	AI038250	711 bp mRNA linear EST 30-JUN-1998											
FEATURES	AI038250	711 bp mRNA linear EST 30-JUN-1998											
source	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query Match	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Best Local Similarity	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Matches	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Query	AI038250	711 bp mRNA linear EST 30-JUN-1998											
Db	AI038250	711 bp mRNA											

California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

ORIGIN

Query Match 32.8%; Score 687.2; DB 2; Length 852;
 Best_Local_Similarity 97.2%; Pred. No. 9.5e-163;
 Matches 785; Conservative 0; Mismatches 13; Indels 10; Gaps 8;

Qy 1 ATGAGCTCCCGAAAGTGTGGCCATTTCAGGCCCGAAAGCGGAGCGCCGAAAAGAGAGAAA 60
 |||||
 Db 15 ATGAGCTCCCGAAAGTGTGGCCATTTCAGGCCCGAAAGCGGAGCGCCGAAAAGAGAGAAA 74
 |||||

Qy 61 CATCCGAAAAGCCGGAGCTTCAACAGAAAGCTCCTTTAGTTCTCTCTCTCCACCGCCA 120
 |||||
 Db 75 CATCCGAAAAGCCGGAGCTTCAACAGAAAGCTCCTTTAGTTCTCTCTCTCCACCGCCA 134
 |||||

Qy 121 CCACCACCCACCGCCACCTTTTGGCAGACCCACACCCCGGAGCCAGAGGAGGATC 180
 |||||
 Db 135 CCAACACACCCACCGCCACCTTTTGGCAGACCCACACCCCGGAGCCAGAGGAGGATC 194
 |||||

Qy 181 CTGGGATCAGATGATGAGGACCAAGAGACCTTCGGGACTACTCCAAAGGTGGATATCAT 240
 |||||
 Db 195 CTGGGATCAGATGATGAGGACCAAGAGACCTTCGGGACTACTCCAAAGGTGGATATCAT 254
 |||||

Qy 241 CCAGTGAAAATTTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGG 300
 |||||
 Db 255 CCAGTGAAAATTTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGG 314
 |||||

Qy 301 GGGCACTTCTACTGTCTGGCTGTCTGGGATATGACGGGAAAAAGATTGTGTGCAATG 360
 |||||
 Db 315 GGGCACTTCTACTGTCTGGCTGTCTGGGATATGACGGGAAAAAGATTGTGTGCAATG 374
 |||||

Qy 361 AAAGTTGTAAAAGTGGCCAGCATTTACGGAGACAGCCTTTGGATGAAATAAAATTCGTC 420
 |||||
 Db 375 AAAGTTGTAAAAGTGGCCAGCATTTACGGAGACAGCCTTTGGATGAAATAAAATTCGTC 434
 |||||

Qy 421 AAATGTGTTCGAGAAAGTGATCCAGATGACCCAAACAAAGACATGGTGGTCCAGCTCAT 480
 |||||
 Db 435 AAATGTGTTCGAGAAAGTGATCCAGATGACCCAAACAAAGACATGGTGGTCCAGCTCAT 494
 |||||

Qy 481 GACGACTTCAAGATTTACGGATGAATGGGATACATGCTGTCATGGTCTTCGAAGTACTT 540
 |||||
 Db 495 GACGACTTCAAGATTTACGGATGAATGGGATACATGCTGTCATGGTCTTCGAAGTACTT 554
 |||||

Qy 541 GGCCACCATCTCCTCAAGTGGATCATCAAAATCCAACTAT - CAAGGCCCTCCCACTAGCTT 598
 |||||
 Db 555 GGCCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATTCAGGGGCTCCCACTAGCTT 614
 |||||

Qy 599 GTG-TGAAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGA-TTACTTACACAGTAAAGT 656
 |||||
 Db 615 GTGTTGAAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTTACTTACACAGTAAAGT 674
 |||||

Qy 657 CAAGATCA-TTCATACCTGA-CATAAAGCCGGAAAAATA-TCTTGATGTGTGGATGATGC 713
 |||||
 Db 675 CAAGATCATTTTCACTACTGACCATAAAGCCGACAAATATCTTCGATGTGTGGATGATGC 734
 |||||

Qy 714 ATATGTCAGAGAGATGCGAGCTGAGGCCACTGAGTGGCAGAAAGCAGGCTGCTCTCTCTCC 773
 |||||
 Db 735 ATATGTCAGAGAGAGAGGGCAG-TGAGGGCCACTGAGTGGCAGAAAGCAGGTG--TCCTCCCC 791
 |||||

Qy 774 TTCAGGGTCTGCAGTGAGTACGGCTCCA 801
 |||||

Db 792 TTCAGGGTCTGCGGTGATTCGGTCCCA 819
 |||||

RESULT 15
 AW149364/c
 LOCUS AW149364
 DEFINITION x336c05.x1 NCI CGAP Brln50 Homo sapiens cDNA clone IMAGE:2620136 3' similar to tr:O75220 O75220 SERINE KINASE SRPK2 ;, mRNA sequence.
 AW149364
 ACCESSION AW149364.1 GI:6197260
 VERSION
 KEYWORDS EST.

```

|||||
Db 313 TCGAAGTACTTGGCCACCATCTCCTCAAGTGGATCATCAATCCAACTATCAAGGCTCC 254
Qy 590 CAGTACGTTGTGTGAAGAGTATCATTCGACAGGTCCTTCAAGGTTAGATTACTTACACA 649
Db 253 CAGTACGTTGTGTAAAGAGTATCATTCGACAGGTCCTTCAAGGTTAGATTACTTACACA 194
Qy 650 GTAAGTGCAAGATCATTTCTACTACTGACATAAAGCCGGAATAATCTTGTGATGTGTGGATG 709
Db 193 GTAAGTGCAAGATCATTTCTACTACTGACATAAAGCCGGAATAATCTTGTGATGTGTGGATG 134
Qy 710 ATGCATATGTGAGAAATGCGAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTC 769
Db 133 ATGCATATGTGAGAAATGCGAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTC 74
Qy 770 CTCCTTCAGGGTCTGCAGTGAGTACCGGCTCCACAGCAGAAACCTATAGGAAAAATATCTA 829
Db 73 CTCCTTCAGGGTCTGCAGTGAGTACCGGCTCCACAGCAGAAACCTATAGGAAAAATATCTA 14
Qy 830 AAAACAAAAGAA 842
Db 13 AAAACAAAAGAA 1
```

Search completed: January 23, 2005, 11:04:31
Job time : 6296.74 secs

This Page Blank (uspio)

181 CTGGATCAGATGATGAGGAGCAAGAGCCCTGCGGACTACTGCAAAAGTGGATATCAT 240
 271 CTGGATCAGATGATGAGGAGCAAGAGCCCTGCGGACTACTGCAAAAGTGGATATCAT 330
 241. CCAAGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAAGCTGGATGG 300
 331 CCAAGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAAGCTGGATGG 390
 301 GGGCACTTCTCTACTGTCTGGCTGTGGGATATGCGAGGGGAAAGATTTGTTGCAATG 360
 391 GGGCACTTCTCTACTGTCTGGCTGTGGGATATGCGAGGGGAAAGATTTGTTGCAATG 450
 361 AAAGTTGTAATAAGTCCAGCAATTTATACGGAGACAGCTTGGATGAATAAATTTGCTC 420
 451 AAAGTTGTAATAAGTCCAGCAATTTATACGGAGACAGCTTGGATGAATAAATTTGCTC 510
 421 AAATGTGTTGCAAAAAGTATCCAGTGAACCCAAACAAAGACATGGTGGTCCAGCTCAT 480
 511 AAATGTGTTGCAAAAAGTATCCAGTGAACCCAAACAAAGACATGGTGGTCCAGCTCAT 570
 481 GACGACTTCAAGATTTCAAGGATGAATGGGATACATGCTGATGGTCTTGAAGTACTT 540
 571 GACGACTTCAAGATTTCAAGGATGAATGGGATACATGCTGATGGTCTTGAAGTACTT 630
 541 GGGCACCATCTCTCAAGTGGATCATCAATCCAACTCAAGGCTCCAGTACGTTGT 600
 631 GGGCACCATCTCTCAAGTGGATCATCAATCCAACTCAAGGCTCCAGTACGTTGT 690
 601 GTGAAGATGATCATTCGACAGGTCTTCAAGGTTAGATTAATTACAGTAAGTGAAG 660
 691 GTGAAGATGATCATTCGACAGGTCTTCAAGGTTAGATTAATTACAGTAAGTGAAG 750
 661 ATCATTCATCTGACATAAAGCGGAAATATCTTGATGTGTGGATGATCATATGTG 720
 751 ATCATTCATCTGACATAAAGCGGAAATATCTTGATGTGTGGATGATCATATGTG 810
 721 AGAAGATGGCACTCAGGCACTGAGTGGCAGAAAGAGGTGCTCTCTCTCTCAGGG 780
 811 AGAAGATGGCACTCAGGCACTGAGTGGCAGAAAGAGGTGCTCTCTCTCTCAGGG 870
 781 TCTGCAATGATGATGCTCCAGCAGAAACCTTATAGGAAATATCTTAAACCAAAAAG 840
 871 TCTGCAATGATGATGCTCCAGCAGAAACCTTATAGGAAATATCTTAAACCAAAAAG 930
 841 AAAAACTGAAAGAAACAGAGAGCGGCTGCTTATGGAGAGCGCTGCGAGGAG 900
 931 AAAAACTGAAAGAAACAGAGAGCGGCTGCTTATGGAGAGCGCTGCGAGGAG 990
 901 ATAGAAGATTTGAGCGGAGAGCTGAAAGGAAATATAGAGAAACATCACTCAGCT 960
 991 ATAGAAGATTTGAGCGGAGAGCTGAAAGGAAATATAGAGAAACATCACTCAGCT 1050
 961 GCACCTTCCAAATGACCGAGATGGGAAATATGCGCCAGAGTGAACTTAAACCAACAGGA 1020
 1051 GCACCTTCCAAATGACCGAGATGGGAAATATGCGCCAGAGTGAACTTAAACCAACAGGA 1110
 1021 TTAGAGAGCGGCTGAGGAGAGAGCTGCAAGGCAATGCTGAAGCTGAGGACCAAGAA 1080
 1111 TTAGAGAGCGGCTGAGGAGAGAGCTGCAAGGCAATGCTGAAGCTGAGGACCAAGAA 1170
 1081 GAGAAAGAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGATCAGGAA 1140
 1171 GAGAAAGAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGATCAGGAA 1230
 1141 CTTGCGAATAGACCTTACGTGATGATCACTTAAACCAATGCGCATTTGAGAT 1200
 1231 CTTGCGAATAGACCTTACGTGATGATCACTTAAACCAATGCGCATTTGAGAT 1290
 1201 GGGCCATTTCTCCTGAGGAGAGCACTGACGATGAAGATGATGATGAAGAGACTGCCCCA 1260
 1291 GGGCCATTTCTCCTGAGGAGAGCACTGACGATGAAGATGATGATGAAGAGACTGCCCCA 1350

1261 AATCTCTGAGGAATATATCTTGATGAGCCAAATGCGAGAAAGTGATTTACATATAGCAGC 1320
 1351 AATCTCTGAGGAATATATCTTGATGAGCCAAATGCGAGAAAGTGATTTACATATAGCAGC 1410
 1321 TCCATGAAACAAATCAATGATGAAATGCGAAATGCGAGCATAAATTTCCGAGTACAG 1380
 1411 TCCATGAAACAAATCAATGATGAAATGCGAAATGCGAGCATAAATTTCCGAGTACAG 1470
 1381 TTCCACAGATTTTCCACCTCGTTGTTCTCTGGATCCCTTAGAACCTGTGGCTCGGGCTCT 1440
 1471 TTCCACAGATTTTCCACCTCGTTGTTCTCTGGATCCCTTAGAACCTGTGGCTCGGGCTCT 1530
 1441 GTGCTTTCTGAGGATCACCATTAAGTACAGCAAGAGAGAGAGTCCATCCATGACAGA 1500
 1531 GTGCTTTCTGAGGATCACCATTAAGTACAGCAAGAGAGAGAGTCCATCCATGACAGA 1590
 1501 AGCAGACGTTTTCAGGCTCCAGTACTGGGATTTGCGAAAGCAAAACCCGGGAGCT 1560
 1591 AGCAGACGTTTTCAGGCTCCAGTACTGGGATTTGCGAAAGCAAAACCCGGGAGCT 1650
 1561 GACTTGTGTGTAATCCCTGGATCCCGGAATGCGATGAGTAAATTTAGAGTAAATTTGCT 1620
 1651 GACTTGTGTGTAATCCCTGGATCCCGGAATGCGATGAGTAAATTTAGAGTAAATTTGCT 1710
 1621 GACTTGTGTGTAATCCCTGGATCCCGGAATGCGATGAGTAAATTTAGAGTAAATTTGCT 1680
 1711 GACTTGTGTGTAATCCCTGGATCCCGGAATGCGATGAGTAAATTTAGAGTAAATTTGCT 1770
 1681 TACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGGAGC 1740
 1771 TACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGGAGC 1830
 1741 ACGCGCTGTATGGCAATTTGAGCTGGCAACCGGAGATTTATTTGTTGAAACCAATTTCTGGG 1800
 1831 ACGCGCTGTATGGCAATTTGAGCTGGCAACCGGAGATTTATTTGTTGAAACCAATTTCTGGG 1890
 1801 GAAAGTATTTCCAGAGCAAGACACATAGCCACATCATAGAGCTGCTAGGCAATTT 1860
 1891 GAAAGTATTTCCAGAGCAAGACACATAGCCACATCATAGAGCTGCTAGGCAATTT 1950
 1861 CCAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTCAATCGCAGAGGAGAA 1920
 1951 CCAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTCAATCGCAGAGGAGAA 2010
 1921 CTGCGACATCATCAAGCTGAAGCTGGAGCTCTTTGATGATCTTTGATGATCTTTGATGAT 1980
 2011 CTGCGACATCATCAAGCTGAAGCTGGAGCTCTTTGATGATCTTTGATGATCTTTGATGAT 2070
 1981 GGCTGGCCCCATGAGATGCTGACAGTTTACAGATTTCCCTGATCCCGATGTTAGAAATG 2040
 2071 GGCTGGCCCCATGAGATGCTGACAGTTTACAGATTTCCCTGATCCCGATGTTAGAAATG 2130
 2041 GTTCCAGAAAAACAGAGCTCAGCTGGCGAATGCTTGGCATCTTGGTGAATTTCT 2097
 2131 GTTCCAGAAAAACAGAGCTCAGCTGGCGAATGCTTGGCATCTTGGTGAATTTCT 2187

RESULT 2

US-10-207-973-1

; Sequence 1. Application US/10207973

; Publication No. US20030175927A1.

; GENERAL INFORMATION:

; APPLICANT: ASU-THREIDEH, Jane et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL001043

; CURRENT APPLICATION NUMBER: US/10/207,973

; CURRENT FILING DATE: 2002-07-31

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 3253

; TYPE: DNA			
; ORGANISM: Human			
US-10-207-973-1			
Query Match		100.0%; Score 2097; DB 15; Length 3253;	
Best Local Similarity		100.0%; Pred. No. 0;	
Matches 2097; Conservative		0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGAGCTCCCGNAAGTGTGGCCATTTCAGGCCGGAAGCGGAGGCGGAAAGAGAGAAA	60
DB	91	ATGAGCTCCCGAAAGTGTGGCCATTTCAGGCCGGAAGCGGAGGCGGAAAGAGAGAAA	150
QY	61	CATCCGAAAAAGCGGAGCCTCAACAGAAAGTCTCTTTAGTTCTCTCCTCCTCCACCGCCA	120
DB	151	CATCCGAAAAAGCGGAGCCTCAACAGAAAGTCTCTTTAGTTCTCTCCTCCTCCACCGCCA	210
QY	121	CCACCAACCAACCGCCACCTTTGGCCAGACCCACACACCCCGGAGCCAGAGGAGATC	180
DB	211	CCACCAACCAACCGCCACCTTTGGCCAGACCCACACACCCCGGAGCCAGAGGAGATC	270
QY	181	CTGGGATCAGATGATGAGGACGAGGACCCCTGGCGACTACTGCGAAGGTGATATCAT	240
DB	271	CTGGGATCAGATGATGAGGACGAGGACCCCTGGCGACTACTGCGAAGGTGATATCAT	330
QY	241	CCAGTGAAATTTGGAGACCTCTTCAATGCGCGTATCATGTTATTAGAAAGCTTGGATG	300
DB	331	CCAGTGAAATTTGGAGACCTCTTCAATGCGCGTATCATGTTATTAGAAAGCTTGGATG	390
QY	301	GGGCACTTCTCTACTGTCTGGCTGTGCTGGGATATGAGGGGAAAAAGATTGTTGCAATG	360
DB	391	GGGCACTTCTCTACTGTCTGGCTGTGCTGGGATATGAGGGGAAAAAGATTGTTGCAATG	450
QY	361	AAAGTTGTAAGTCCCGACGATTTACGGAGACGCTTGGATGGAATGAAATAAAATGCTC	420
DB	451	AAAGTTGTAAGTCCCGACGATTTACGGAGACGCTTGGATGGAATGAAATAAAATGCTC	510
QY	421	AAATGTGTCGAGAAAGTGCATCCAGTGACCCAAACAAAGACATGCTGCTCCAGCTCAT	480
DB	511	AAATGTGTCGAGAAAGTGCATCCAGTGACCCAAACAAAGACATGCTGCTCCAGCTCAT	570
QY	481	GAGCACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGTCATGCTCTTCGAAGTACT	540
DB	571	GAGCACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGTCATGCTCTTCGAAGTACT	630
QY	541	GGCCACCATCTCTCAAGTGGATCATCAATCCAACTATCAAGGCCCTCCAGGTAGTGT	600
DB	631	GGCCACCATCTCTCAAGTGGATCATCAATCCAACTATCAAGGCCCTCCAGGTAGTGT	690
QY	601	GTGAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGCAG	660
DB	691	GTGAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGCAG	750
QY	661	ATCATTATCATCTGACATAAAGCCGGAATAATCTTGATGTGTGGATGATGATATGTG	720
DB	751	ATCATTATCATCTGACATAAAGCCGGAATAATCTTGATGTGTGGATGATGATATGTG	810
QY	721	AGAAGAATGGCAGTGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCCTCCTCAGGG	780
DB	811	AGAAGAATGGCAGTGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCCTCCTCAGGG	870
QY	781	TCTGCAGTATGATCGGCTCCACAGCAGAACCTTATAGGAAATAATCTTAAACAACAAA	840
DB	871	TCTGCAGTATGATCGGCTCCACAGCAGAACCTTATAGGAAATAATCTTAAACAACAAA	930
QY	841	AAAAAACTGAAAAAGAAAACAGAGAGCGCAGGTGAGTTATTGGAGAGCGCCTGCAAGG	900
DB	931	AAAAAACTGAAAAAGAAAACAGAGAGCGCAGGTGAGTTATTGGAGAGCGCCTGCAAGG	990
QY	901	ATAGAAAGATTGGAGCGAGAGCTGAAAGGAAAAATTAATAGAGAAAAATCACTCCTCAG	960
DB	991	ATAGAAAGATTGGAGCGAGAGCTGAAAGGAAAAATTAATAGAGAAAAATCACTCCTCAG	1050
QY	961	GCACCTTCCATGACCAGGATGGCGAATACTGCCAGAGGTGAAACTAAACACACAGGA	1020

1051	GCACCTTCCATGACCGAGATGGCGAATACTGCCAGAGGTGAAACTTAAACACACAGGA	1110
1021	TTAGAGGAGCGGCTGAGCGACAGACTGCAAAAGGCAATTTGGTGAAGCTGAGGACCAAGAA	1080
1111	TTAGAGGAGCGGCTGAGCGACAGACTGCAAAAGGCAATTTGGTGAAGCTGAGGACCAAGAA	1170
1081	GAGAAAGAGATGCTGAGNAAGAAACATTTGAAAAGATGAAAGTATGATCTAGATCAGGA	1140
1171	GAGAAAGAGATGCTGAGNAAGAAACATTTGAAAAGATGAAAGTATGATCTAGATCAGGA	1230
1141	CTTGGCAACATAGACCTACGTGGATAGAAATCACTTAAACCAATGCCCATATTGAGAAT	1200
1231	CTTGGCAACATAGACCTACGTGGATAGAAATCACTTAAACCAATGCCCATATTGAGAAT	1290
1201	GGCCCATCTCTACTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAAGCTGCCCA	1260
1291	GGCCCATCTCTACTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAAGCTGCCCA	1350
1261	ANTCTGTAGGAATATATCTTGATGAGCCAAATGCAAGATGATGATGATGATGATGATGAT	1320
1351	ANTCTGTAGGAATATATCTTGATGAGCCAAATGCAAGATGATGATGATGATGATGATGAT	1410
1321	TCCTATGAACAAATTCATGCTGAAATTCCTGATGAGCCAAATGCAAGATGATGATGATGAT	1380
1411	TCCTATGAACAAATTCATGCTGAAATTCCTGATGAGCCAAATGCAAGATGATGATGATGAT	1470
1381	TTCCAGAGTTTCCACCTCGTTGTTCTCTGATCCTTAGAACCTGTGTGCTGCGCTCT	1440
1471	TTCCAGAGTTTTCACCTCGTTGTTCTCTGATCCTTAGAACCTGTGTGCTGCGCTCT	1530
1441	GTGCTTTCTGAGGATCACCCTTACTGAGCAAGGAGGAGAGCAGTCCATCCCATGACAGA	1500
1531	GTGCTTTCTGAGGATCACCCTTACTGAGCAAGGAGGAGAGCAGTCCATCCCATGACAGA	1590
1501	ACGAGAACGGTTTTCAGCTCAGTACTGCGGATTTGCCAAAAGCAAAACCCGGGAGCT	1560
1591	ACGAGAACGGTTTTCAGCTCAGTACTGCGGATTTGCCAAAAGCAAAACCCGGGAGCT	1650
1561	GACTTGTGTTGTAATCCCTGAGTCCGCGAATGCAAGATGATGATGATGATGATGATGAT	1620
1651	GACTTGTGTTGTAATCCCTGAGTCCGCGAATGCAAGATGATGATGATGATGATGATGAT	1710
1621	GACTTGTGTTGTAATCCCTGAGTCCGCGAATGCAAGATGATGATGATGATGATGATGAT	1680
1711	GACTTGTGTTGTAATCCCTGAGTCCGCGAATGCAAGATGATGATGATGATGATGATGAT	1770
1681	TACCGCTCCATAGAGTCTTAAATAGAGCGGGTACAGCACCCCTGCGACATCTCGAGC	1740
1771	TACCGCTCCATAGAGTCTTAAATAGAGCGGGTACAGCACCCCTGCGACATCTCGAGC	1830
1741	ACGGGCTGTATGGCATTTGAGCTGGCAACGGGAGATTTATTTGTTGAAACCAATTTCTGG	1800
1831	ACGGGCTGTATGGCATTTGAGCTGGCAACGGGAGATTTATTTGTTGAAACCAATTTCTGG	1890
1801	GAGACTATTTCAGAGACGAAAGACCATAGCCCATCATATAGAGTGTCTAGCAGTATT	1860
1891	GAGACTATTTCAGAGACGAAAGACCATAGCCCATCATATAGAGTGTCTAGCAGTATT	1950
1861	CCAAGCACTTTGCTCTATCTGAAAAATTTCTCGGAAATTTCTTCAATCGCAGAGAGAA	1920
1951	CCAAGCACTTTGCTCTATCTGAAAAATTTCTCGGAAATTTCTTCAATCGCAGAGAGAA	2010
1921	CTGCGCACATCAACCAAGCTGAAGCCCTGGAGCCTCTTTGATGATGATGATGATGATGAT	1980
2011	CTGCGCACATCAACCAAGCTGAAGCCCTGGAGCCTCTTTGATGATGATGATGATGATGAT	2070
1981	GGCTGGCCCCATGAGATGCTGCAAGTTTACAGATTTCTGATCCCGATGTTAGAAATG	2040
2071	GGCTGGCCCCATGAGATGCTGCAAGTTTACAGATTTCTGATCCCGATGTTAGAAATG	2130
2041	GTTCCAGAAAAAGCGCTCAGCTGGCGAATGCCCTCGGCATCCTTGTGATTTCT	2097

Db 2131 GTTCCAGAAAACGAGCCTCAGCTGGCGAATGCTTGGCATCCTTGGTGAATTC 2187

RESULT 3

US-10-799-676-1

; Sequence 1, Application US/10799676

; Publication No. US20040157297A1

; GENERAL INFORMATION:

; APPLICANT: ARU-THREIDEH, Jane et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CL001043DIV II

; CURRENT APPLICATION NUMBER: US/10799,676

; PRIORITY FILING DATE: 2004-03-15

; PRIOR APPLICATION NUMBER: 10/207,973

; PRIOR FILING DATE: 2002-07-31

; PRIOR APPLICATION NUMBER: 09/759,359

; PRIOR FILING DATE: 2001-01-16

; NUMBER OF SEQ ID NOS: 3

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 3253

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-799-676-1

Query Match 100.0%; Score 2097; DB 17; Length 3253;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCTCCCGAAAGTGTGCTGCTTCAAGGCCGAGAGCGGAGCGGAGAGAGAGAAA 60

Db 91 ATGAGCTCCCGAAAGTGTGCTGCTTCAAGGCCGAGAGCGGAGCGGAGAGAGAGAAA 150

QY 61 CATCCGAAAAAGCGGAGCGCTCAACAGAAAAGTCTCTTTAGTTCTCTCTCCACCGCCA 120

Db 151 CATCCGAAAAAGCGGAGCGCTCAACAGAAAAGTCTCTTTAGTTCTCTCTCCACCGCCA 210

QY 121 CCACACACACCGCGCCCTTTGCCAGACCCACACCCCGGAGCGGAGAGAGAGATC 180

Db 211 CCACACACACCGCGCCCTTTGCCAGACCCACACCCCGGAGCGGAGAGAGATC 270

QY 181 CTGGATCAGATGATGAGGAGCAAGAGACCTCTGCGGACTACTGCAAAAGGTGGATATCAT 240

Db 271 CTGGATCAGATGATGAGGAGCAAGAGACCTCTGCGGACTACTGCAAAAGGTGGATATCAT 330

QY 241 CCAGTGAAAATTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAAGCTTGGATGG 300

Db 331 CCAGTGAAAATTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAAGCTTGGATGG 390

QY 301 GGGCACTTCTTACTCTGCTGGCTGTCTGGATATGCGGGGAAAAGATTGTTGCAATG 360

Db 391 GGGCACTTCTTACTCTGCTGGCTGTCTGGATATGCGGGGAAAAGATTGTTGCAATG 450

QY 361 AAAAGTTGAAAAGTCCCGAGCAATTATACGGAGACAGCCCTTGGATGAAAATAAAATTGCTC 420

Db 451 AAAAGTTGAAAAGTCCCGAGCAATTATACGGAGACAGCCCTTGGATGAAAATAAAATTGCTC 510

QY 421 AAATGTTGCGAAAAGTATCCAGTGCACCCCAAAAGACATGCTGCTCCAGCTCATTT 480

Db 511 AAATGTTGCGAAAAGTATCCAGTGCACCCCAAAAGACATGCTGCTGCTCCAGCTCATTT 570

QY 481 GAGCACTTCAAGATTTTCAAGCATGAATGGATACATGCTGCAATGCTTCTGCAAGTACTT 540

Db 571 GAGCACTTCAAGATTTTCAAGCATGAATGGATACATGCTGCAATGCTTCTGCAAGTACTT 630

QY 541 GGCACCATCTCTCAAGTGGATCATCAAACTCAACTCAAGGCTCCAGCTGACGTTGT 600

Db 631 GGCACCATCTCTCAAGTGGATCATCAAACTCAACTCAAGGCTCCAGTACGTTGT 690

QY 601 GTCAAGATGATCATTCGACAGGTCCTTCAAGGTTAGATTACTTACAGATGATGCAAG 660

Db 691 GTGAAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTACTTTACACAGTAAGTCAAG 750

QY 661 ATCAATTCATCTGACATAAAAGCCGAAAAATATCTTTGATGTGTGGATGATGATATGTG 720

Db 751 ATCAATTCATCTGACATAAAAGCCGAAAAATATCTTTGATGTGTGGATGATGATATGTG 810

QY 721 AGAAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAAGCAGGTGCTCTCTCTTCCAGGG 780

Db 811 AGAAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAAGCAGGTGCTCTCTCTTCCAGGG 870

QY 781 TCTGAGTGTAGTACGGCTCCACACAGAAAACCTATAGAAAAATATCTAAAAACAAAAAG 840

Db 871 TCTGAGTGTAGTACGGCTCCACACAGAAAACCTATAGAAAAATATCTAAAAACAAAAAG 930

QY 841 AAAAACTGAAAAAGAAAACAGAGAGGCGAGCTCAGTTATTGGAGAAAGCGCTCGAGAG 900

Db 931 AAAAACTGAAAAAGAAAACAGAGAGGCGAGCTCAGTTATTGGAGAAAGCGCTCGAGAG 990

QY 901 ATAGAAGAATTTGGAGCGAGAAAGCTGAAAGGAAAAATAATAGAAAAACATCACCTCAGCT 960

Db 991 ATAGAAGAATTTGGAGCGAGAAAGCTGAAAGGAAAAATAATAGAAAAACATCACCTCAGCT 1050

QY 961 GCACCTTCCATGACACGAGTGGCGAATACCTGCCAGAGGTGAAACTAAAAACAAACAGA 1020

Db 1051 GCACCTTCCATGACACGAGTGGCGAATACCTGCCAGAGGTGAAACTAAAAACAAACAGA 1110

QY 1021 TTAGAGGAGCGCGCTGAGGCGAGAGACTGCCAAGGACAAATGCTGAAGCTGAGGACCAAGAA 1080

Db 1111 TTAGAGGAGCGCGCTGAGGCGAGAGACTGCCAAGGACAAATGCTGAAGCTGAGGACCAAGAA 1170

QY 1081 GAGAAAGAAGATGCTGAGAAAAAACAATTGAAAAAGATGAAGATGATGTAGATCAGGAA 1140

Db 1171 GAGAAAGAAGATGCTGAGAAAAAACAATTGAAAAAGATGAAGATGATGTAGATCAGGAA 1230

QY 1141 CTTCGGAACATAGACCTTACCTGATGAAATCACTAAACCAATGGCCAATTATTGAGAAAT 1200

Db 1231 CTTCGGAACATAGACCTTACCTGATGAAATCACTAAACCAATGGCCAATTATTGAGAAAT 1290

QY 1201 GGCCCATCTCTCACTGGAGCAGCAACTGGACGATGAAGATGATGAAGAGAGACTGCCCA 1260

Db 1291 GGCCCATCTCTCACTGGAGCAGCAACTGGACGATGAAGATGATGAAGAGAGACTGCCCA 1350

QY 1261 AATCTGAGGAATATATCTTTGATGAGCAAAATGCAAGAAAGTGAATACATATAGCAGC 1320

Db 1351 AATCTGAGGAATATATCTTTGATGAGCAAAATGCAAGAAAGTGAATACATATAGCAGC 1410

QY 1321 TCCTATGAACAAATTCATGTTGAAATGGCAATGGACGACATAAATTTCCGAGTCAAG 1380

Db 1411 TCCTATGAACAAATTCATGTTGAAATGGCAATGGACGACATAAATTTCCGAGTCAAG 1470

QY 1381 TTCCAGAGTTTTCACCTCGTTGTTCTCTGGATCTTTAGAACCTGTGGCTGCGGCTCT 1440

Db 1471 TTCCAGAGTTTTCACCTCGTTGTTCTCTGGATCTTTAGAACCTGTGGCTGCGGCTCT 1530

QY 1441 GTGCTTTCTGAGGATACCACTTACTGAGCAAGGAGAGAGAGTCCATCCCATCAGAGA 1500

Db 1531 GTGCTTTCTGAGGATACCACTTACTGAGCAAGGAGAGAGAGTCCATCCCATCAGAGA 1590

QY 1501 AGCAGAACGGTTTTCAGCTTCCAGTCTGCGGATTTTCCAAAAAGCAAAAAACCCGGGAGCT 1560

Db 1591 AGCAGAACGGTTTTCAGCTTCCAGTCTGCGGATTTTCCAAAAAGCAAAAAACCCGGGAGCT 1650

QY 1561 GACTTTGTTGGTGAATTCCTCGATCCGCGGAATGAGATAAAATTTAGAGTAAAAATTGCT 1620

Db 1651 GACTTTGTTGGTGAATTCCTCGATCCGCGGAATGAGATAAAATTTAGAGTAAAAATTGCT 1710

QY 1621 GACCTGGGAAATCTTTGTTGGTGCATAAACACTTTCAGGAGACATCCAGACGGCTCAG 1680

Db 1711 GACCTGGGAAATCTTTGTTGGTGCATAAACACTTTCAGGAGACATCCAGACGGCTCAG 1770

QY 1681 TACCCCTCCATAGAGTTTAAATAGGAGCGGGGTACAGACCCCTTCGCGACATCTGGAGC 1740

Db 1771 TACCCCTCCATAGAGTTTAAATAGGAGCGGGGTACAGACCCCTTCGCGACATCTGGAGC 1830

Qy	1741	ACGGCGTGATGGCAATTTGACTGGCAACGGGAGATTATTTGTTTGAACACATTTCTGGG	1800
Db	1831	ACGGCGTGATGGCAATTTGACTGGCAACGGGAGATTATTTGTTTGAACACATTTCTGGG	1890
Qy	1801	GAAGACTATTCCACAGACGAGACCAATAGCCCCACATCATAGAGCTGCTAGGCAGTATT	1860
Db	1891	GAAGACTATTCCACAGACGAGACCAATAGCCCCACATCATAGAGCTGCTAGGCAGTATT	1950
Qy	1861	CCAAGGCACTTTGCTCTATCTGGAAATATTCTCGGAAATTTCTTCAATCGCAGAGGAA	1920
Db	1951	CCAAGGCACTTTGCTCTATCTGGAAATATTCTCGGAAATTTCTTCAATCGCAGAGGAA	2010
Qy	1921	CTGCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTGTATGTACTTGTGGAAAAAGTAT	1980
Db	2011	CTGCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTGTATGTACTTGTGGAAAAAGTAT	2070
Qy	1981	GGCTGGGCCCAATGAAGATGCTGCACAGTTTACAGATTTCCTGATCCCGATGTTAGAAATG	2040
Db	2071	GGCTGGGCCCAATGAAGATGCTGCACAGTTTACAGATTTCCTGATCCCGATGTTAGAAATG	2130
Qy	2041	GTTTCCAGAAAAACGAGCCTGAGCTGGCGAATGCCCTTCGGCATCCTTGTTGAATTCT	2097
Db	2131	GTTTCCAGAAAAACGAGCCTGAGCTGGCGAATGCCCTTCGGCATCCTTGTTGAATTCT	2187

RESULT 4

US-10-618-941-36 ; Sequence 36, Application US/10618941 ; Publication No. US20040197792A1

```

% GENERATED INFORMATION:
% APPLICANT: WHYTE, DAVID
% APPLICANT: MANNING, GERARD
% APPLICANT: CAENEPEEL, SEAN
% TITLE OF INVENTION: NOVEL KINASES
% FILE REFERENCE: 034536-0321
% CURRENT APPLICATION NUMBER: US/10/618,941
% CURRENT FILING DATE: 2003-07-15
% PRIOR APPLICATION NUMBER: 60/395,632
% PRIOR FILING DATE: 2002-07-15
% NUMBER OF SEQ ID NOS: 143
% SOFTWARE: Patent in version 3.2
% SEQ ID NO 36
% LENGTH: 3715
% TYPE: DNA
% ORGANISM: Homo sapiens
US-10-618-941-36

```

Query Match	96.7%	Score	2028.6	DB 18	Length	3715
Best Local Similarity	99.8%	Pred. No.	0			
Matches 2031	Conservative	0	Mismatches	4	Indels	0
Qy	63	TCCGAAAAAGCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCTCCACGCCACC	122			
Db	208	TTCAGAAAGGCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCTCCACGCCACC	267			
Qy	123	ACCACCACCGCCACCTTTGGCCAGACCCACACCCCGGAGCCAGAGGAGGAGATCCT	182			
Db	268	ACCACCACCGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGGAGATCCT	327			
Qy	183	GGGATCAGATGATGAGGAGCAAGAGGACCTTCGGACTATCTGCAAAAGGTGGATATCATCC	242			
Db	328	GGGATCAGATGATGAGGAGCAAGAGGACCTTCGGACTACTGCAAAAGGTGGATATCATCC	387			
Qy	243	AGTGAAAATTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGGG	302			
Db	388	AGTGAAAATTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGGG	447			
Qy	303	GCATTTCTTACTGTCTGGCTGTCTGGGATATGCAGGGGAAAAAGATTGTTGCAATGAA	362			
Db	448	GCATTTCTTACTGTCTGGCTGTCTGGGATATGCAGGGGAAAAAGATTGTTGCAATGAA	507			
Qy	363	AGTTGTAAAAAGTGCACGACATTATACGGAGACAGCCTTGGATGAAATTAATAATTTGCTCAA	422			

508	Db		AGTTGTAANAAGTGCCAGCATTATACGSGAGACAGCCTTGATGAANAATAAATTTGCTCAA	567
423	Qy		ATGTGTTTCGAGAAAGTGATCCCAAGTACCACAAACAAAGACATGGTGGTCCAGCTCATTTGA	482
568	Db		ATGTGTTTCGAGAAAGTGATCCCAAGTACCACCAACAAAGACATGGTGGTCCAGCTCATTTGA	627
483	Qy		CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCATGGTCTTTCGAAAGTACTTGG	542
628	Db		CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCATGGTCTTTCGAAAGTACTTGG	687
543	Qy		CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCTCCCACTAGTGTCTGT	602
688	Db		CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCTCCCACTAGTGTCTGT	747
603	Qy		GAAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGCAGAT	662
748	Db		GAAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGCAGAT	807
663	Qy		CATTTCATATCGACATAAAGCCGAAATAATCTTGAATGTGTGGATGATCATATGTGAG	722
808	Db		CATTTCATATCGACATAAAGCCGAAATAATCTTGAATGTGTGGATGATCATATGTGAG	867
723	Qy		AAGNATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCCTCCTTCAGGGTC	782
868	Db		AAGNATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCCTCCTTCAGGGTC	927
783	Qy		TGCAGTCAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTAAAAACAAAAAGAA	842
928	Db		TGCAGTCAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTAAAAACAAAAAGAA	987
843	Qy		AAAACTGAAAAAGAAAACAGAAAGGCGAGGCTGAGTTATTGGAGAAGCGCCTGCAGGAGAT	902
988	Db		AAAACTGAAAAAGAAAACAGAAAGGCGAGGCTGAGTTATTGGAGAAGCGCCTGCAGGAGAT	1047
903	Qy		AGAAGAAATTCGAGCGGAGAACTGAAGGAAAAATTAATAGAGAAAAACATCACTCAGCTGC	962
1048	Db		AGAAGAAATTCGAGCGGAGAACTGAAGGAAAAATTAATAGAGAAAAACATCACTCAGCTGC	1107
963	Qy		ACCTTCCAAATGACACAGGATGCGAAATACTGCCACAGGTCGAAACTAAAAACACAGGATT	1022
1108	Db		ACCTTCCAAATGACACAGGATGCGAAATACTGCCACAGGTCGAAACTAAAAACACAGGATT	1167
1023	Qy		AGAGGAGCGGCTGAGSCAGAGACTGCAAAAGGACAAATGGTGAAGCTGAGGACAGGAAGA	1082
1168	Db		AGAGGAGCGGCTGAGSCAGAGACTGCAAAAGGACAAATGGTGAAGCTGAGGACAGGAAGA	1227
1083	Qy		GAAGAAGATGCTGAGAAAGAAAAACATTGAAAAAGATGAAGATGATGTAGATCAGGNACT	1142
1228	Db		GAAGAAGATGCTGAGAAAGAAAAACATTGAAAAAGATGAAGATGATGTAGATCAGGNACT	1287
1143	Qy		TGCGAATATAGACCCCTACGTGGATAGAAATCAACTAAAAACCAATGGCCATATTCAGAAATGG	1202
1288	Db		TGCGAATATAGACCCCTACGTGGATAGAAATCAACTAAAAACCAATGGCCATATTCAGAAATGG	1347
1203	Qy		CCCATTCCTACTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAAGACTGCCCCAAA	1262
1348	Db		CCCATTCCTACTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAAGACTGCCCCAAA	1407
1263	Qy		TCCTGAGGAATATAATCTTGATGAGCCAAATGAGAAAGTGAATTAACATATAGCAGCTC	1322
1408	Db		TCCTGAGGAATATAATCTTGATGAGCCAAATGAGAAAGTGAATTAACATATAGCAGCTC	1467
1323	Qy		CTATGAACAATTCAAATGGTGAATTTGCCAATGACGACATATAAATTTCCCGAGTCACAGTT	1382
1468	Db		CTATGAACAATTCAAATGGTGAATTTGCCAATGACGACATATAAATTTCCCGAGTCACAGTT	1527
1383	Qy		CCGAGAGTTTTCCACCTCGTGTGTTCTCTGGATCCTTAGAACCTGTGGCCTCGCGCTCTGT	1442
1528	Db		CCGAGAGTTTTCCACCTCGTGTGTTCTCTGGATCCTTAGAACCTGTGGCCTCGCGCTCTGT	1587
1443	Qy		GCTTTCTGAGGGATACCACTTACTGAGCAAGGAGGACGAGTCCATCCCATGACAGAAG	1502

Db 1588 GCTTTCTCAGGGATCACCACCTTACTGTGCAAGAGGAGGAGCAGTCCATCCCATGACAGAAG 1647
Qy 1503 CAGAACGGTTTCAGCCTCAGTACTGCGGATTTGCGCAAGCAAAACCCCGGCGAGCTGA 1562
Db 1648 CAGAACGGTTTCAGCCTCAGTACTGCGGATTTGCGCAAGCAAAACCCCGGCGAGCTGA 1707
Qy 1563 CTGTTGTTGTAATCCCTCGGATCCGCGAATGCGAATAAATAGATATAAATTTGCTGA 1622
Db 1708 CTGTTGTTGTAATCCCTCGGATCCGCGAATGCGAATAAATAGATATAAATTTGCTGA 1767
Qy 1623 CTTGGGAAATGCTTCTGCGTGCATATAACATCTTACGGAAGACATCCAGACGCGTCACTA 1682
Db 1768 CTTGGGAAATGCTTCTGCGTGCATATAACATCTTACGGAAGACATCCAGACGCGTCACTA 1827
Qy 1683 CCGCTCCATAGAGTTTAAATAGGAGCGGGGTACAGCACCCCTGCGGACATCTCGAGCAC 1742
Db 1828 CCGCTCCATAGAGTTTAAATAGGAGCGGGGTACAGCACCCCTGCGGACATCTCGAGCAC 1887
Qy 1743 GCGCTGTATGGCATTTGAGCTGGCAACGGGAGATTATTTGTTGAACACACATTTCTGGGGA 1802
Db 1888 GCGGTGTATGGCATTTGAGCTGGCAACGGGAGATTATTTGTTGAACACACATTTCTGGGGA 1947
Qy 1803 AGACTATTCAGAGACGAAGACACATAGCCACATCATAGAGCTGCTAGGAGATATTC 1862
Db 1948 AGACTATTCAGAGACGAAGACACATAGCCACATCATAGAGCTGCTAGGAGATATTC 2007
Qy 1863 AAGGCATTTGCTCTATCTGGAATAATTTCTCGGAATTTCTCAATCGCAGAGGAACT 1922
Db 2008 AAGGCATTTGCTCTATCTGGAATAATTTCTCGGAATTTCTCAATCGCAGAGGAACT 2067
Qy 1923 GCGACACATCACAAGCTGAAGCCCTGGAGCCTCTTTGATGACTTTGTTGGAAGATATGG 1982
Db 2068 GCGACACATCACAAGCTGAAGCCCTGGAGCCTCTTTGATGACTTTGTTGGAAGATATGG 2127
Qy 1983 CTGGCCCATGAGATGCTGCAAGTTTACAGATTTACAGATTTCTCGATCCCGATGTTAGAAATGGT 2042
Db 2128 CTGGCCCATGAGATGCTGCAAGTTTACAGATTTTCTGATCCCGATGTTAGAAATGGT 2187
Qy 2043 TCCAGAAAACGAGCCTCAGCTGGCAATGCGCTTCGGCATCTTGGTTGAATTC 2097
Db 2188 TCCAGAAAACGAGCCTCAGCTGGCAATGCGCTTCGGCATCTTGGTTGAATTC 2242

RESULT 5
US-10-198-846-10286
; Sequence 10286, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 10286
; LENGTH: 4698
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-10286
Query Match 96.7%; Score 2028.6; DB 14; Length 4698;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2031; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Qy 63 TCCGAAAAAGCCGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACCGGCACC 122

Db 254 TTCAAGAAAGCCGAGCCTCAACAGAAAGCTCTCTTTAGTTCTCTCTCCACCGCAC 313
Qy 123 ACCACCAACACCCGAGCCTTTGCGAGACCCACACCCCGGAGCCAGAGGAGAGATCCT 182
Db 314 ACCACCAACACCCGAGCCTTTGCGAGACCCACACCCCGGAGCCAGAGGAGAGATCCT 373
Qy 183 GGGATCAGATGATGAGGAGCAAGAGACCCCTGCGGACTACTGCAAAAGGTGGATATCATCC 242
Db 374 GGGATCAGATGATGAGGAGCAAGAGACCCCTGCGGACTACTGCAAAAGGTGGATATCATCC 433
Qy 243 AGTGAAATTTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTCGATGGGG 302
Db 434 AGTGAAATTTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTCGATGGGG 493
Qy 303 GCATCTCTACTGCTGCGGCTGTGCGGATATGCGAGGGAAGAAAGATTTTTCGAATGAA 362
Db 494 GCATCTCTACTGCTGCGGCTGTGCGGATATGCGAGGGAAGAAAGATTTTTCGAATGAA 553
Qy 363 AGTTGTAAAAAGTCCCGAGCATTTATACGAGACAGAGCCTTGGATGAAATAAATTTGCTCAA 422
Db 554 AGTTGTAAAAAGTCCCGAGCATTTATACGAGACAGAGCCTTGGATGAAATAAATTTGCTCAA 613
Qy 423 ATGTTTGCAGAAAGTATCCAGTGACCCAAACAAAGACATGCTGCTCAGCTCATGA 482
Db 614 ATGTTTGCAGAAAGTATCCAGTGACCCAAACAAAGACATGCTGCTCAGCTCATGA 673
Qy 483 CGATCTCAAGATTTTCAAGGCTGATGCGGATATGCTGATGCTTTCGAAGTACTTGG 542
Db 674 CGATCTCAAGATTTTCAAGGCTGATGCGGATATGCTGATGCTTTCGAAGTACTTGG 733
Qy 543 CCACCATCTCTCAAGTGGATCATCAATCCAACTCAAGGCTCCAGGATGCTTGTGT 602
Db 734 CCACCATCTCTCAAGTGGATCATCAATCCAACTCAAGGCTCCAGGATGCTTGTGT 793
Qy 603 GAAAGTATCAATTTGCGAGGCTCTTCAAGGCTTATGATTTACTTACACAGTAAAGTCAAGAT 662
Db 794 GAAAGTATCAATTTGCGAGGCTCTTCAAGGCTTATGATTTACTTACACAGTAAAGTCAAGAT 853
Qy 663 CATTCATACTGACATAAAGCCGGAATAATCTTCTGATGCTGCGATGATGATGAG 722
Db 854 CATTCATACTGACATAAAGCCGGAATAATCTTCTGATGCTGCGATGATGATGAG 913
Qy 723 AAGAAATGCGAGCTGAGGCTGAGTGGCAGAAAGCAGGCTCTCTCTCTCAGGGTC 782
Db 914 AAGAAATGCGAGCTGAGGCTGAGTGGCAGAAAGCAGGCTCTCTCTCTCAGGGTC 973
Qy 783 TGCAGTGTACGCTCCACAGCAAAACCTATAGGAAAAATATCTTAAAAACAAAAAGAA 842
Db 974 TGCAGTGTACGCTCCACAGCAAAACCTATAGGAAAAATATCTTAAAAACAAAAAGAA 1033
Qy 843 AAAAATGAAAAAGAAACAGAGCGGCTGAGTATTTGGAAGAGCGCCTGCGAGGAT 902
Db 1034 AAAAATGAAAAAGAAACAGAGCGGCTGAGTATTTGGAAGAGCGCCTGCGAGGAT 1093
Qy 903 AGAAGATTTGAGCGAGAGCTGAAAGAAATAATAGAAAGAAACATCAGCTCAGCTGC 962
Db 1094 AGAAGATTTGAGCGAGAGCTGAAAGAAATAATAGAAAGAAACATCAGCTCAGCTGC 1153
Qy 963 ACCTTCCCAATGACAGGATGCGAATCTGCCAGAGGTGAACTTAAAAACAAACAGGAT 1022
Db 1154 ACCTTCCCAATGACAGGATGCGAATCTGCCAGAGGTGAACTTAAAAACAAACAGGAT 1213
Qy 1023 AGAGAGGCGGCTGAGGCGAGACATGCAAGAGCAATGCTGAGCTGAGGACCGAGAGA 1082
Db 1214 AGAGAGGCGGCTGAGGCGAGACATGCAAGAGCAATGCTGAGCTGAGGACCGAGAGA 1273
Qy 1083 GAAAGAGATCTGAGAAAGAAACATTTGAAAGATGATGATGATGATGATGATGATGATGAT 1142
Db 1274 GAAAGAGATCTGAGAAAGAAACATTTGAAAGATGATGATGATGATGATGATGATGATGAT 1333
Qy 1143 TCCGAAACATAGACCCCTACGTTGATAGAAATCACCTTAAAAACCAATGCCCATATTTGAGATGG 1202

Db 1334 TCGAATCATAGACCCCTACGTGGATAGATCACTCAAAACCAATGGCCATATTGAGAATGG 1393
Qy 1203 CCCATTCTCACTGGACGACCACTGACGATGAAGATGATGAAGAAGATGCCCAAA 1262
Db 1394 CCCATTCTCACTGGACGACCACTGACGATGAAGATGATGAAGAAGATGCCCAAA 1453
Qy 1263 TCTGAGGATATAATCTTGTATGAGCCAAATGAGAAAGTGAATTACACATATAGCAGCTC 1322
Db 1454 TCTGAGGATATAATCTTGTATGAGCCAAATGAGAAAGTGAATTACACATATAGCAGCTC 1513
Qy 1323 CTATGAACAAATCAATGGTGAATTTGCCAAATGACACATATAAAATTTCCGAGTCAAGTT 1382
Db 1514 CTATGAACAAATCAATGGTGAATTTGCCAAATGACACATATAAAATTTCCGAGTCAAGTT 1573
Qy 1383 CCCAGATTTTCCACCTCGTTCTCTGATCCCTTAGAACCTGTGGCTCGGGCTCTGT 1442
Db 1574 CCCAGATTTTCCACCTCGTTCTCTGATCCCTTAGAACCTGTGGCTCGGGCTCTGT 1633
Qy 1443 GCTTTCTGAGGATCAACCTTACTGAGCAAGAGGAGAGCAGTCCATCCCATGACAGAAG 1502
Db 1634 GCTTTCTGAGGATCAACCTTACTGAGCAAGAGGAGAGCAGTCCATCCCATGACAGAAG 1693
Qy 1503 CAGAACGGTTTCAAGCTTCCAGTACTCGGGATTTGCCAAAGCAAAACCCGGGACGCTGA 1562
Db 1694 CAGAACGGTTTCAAGCTTCCAGTACTCGGGATTTGCCAAAGCAAAACCCGGGACGCTGA 1753
Qy 1563 CTTGTTGGTGAATCCCTCGATCCCGGGAATGACAGATAAAATTAGAGTAAATTTCTCTGA 1622
Db 1754 CTTGTTGGTGAATCCCTCGATCCCGGGAATGACAGATAAAATTAGAGTAAATTTCTCTGA 1813
Qy 1623 CTTGGAATGCTTTGTTGGTGCATAAACAATTACGGAGAGACATCCAGACGGTCACTA 1682
Db 1814 CTTGGAATGCTTTGTTGGTGCATAAACAATTACGGAGAGACATCCAGACGGTCACTA 1873
Qy 1683 CCGCTCCATAGAGTTTAAATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAGCAC 1742
Db 1874 CCGCTCCATAGAGTTTAAATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAGCAC 1933
Qy 1743 GCGTGTATGGCAATTTGAGCTGCGCAACGGGAGATTTTGTGTAACCAATTTCTGGGA 1802
Db 1934 GCGTGTATGGCAATTTGAGCTGCGCAACGGGAGATTTTGTGTAACCAATTTCTGGGA 1993
Qy 1803 AGACTATCCAGAGAGAGACACATAGCCACATAGAGCTCTAGGACGATTTCC 1862
Db 1994 AGACTATCCAGAGAGAGACACATAGCCACATAGAGCTCTAGGACGATTTCC 2053
Qy 1863 AAGGCATTTGTCTATCTGGAAATATTTCTCGGAAATTTCTCAATCGCAGAGGAACT 1922
Db 2054 AAGGCATTTGTCTATCTGGAAATATTTCTCGGAAATTTCTCAATCGCAGAGGAACT 2113
Qy 1923 GCGACATCAACAAGCTGAAGCCCTGGAGCTCTTTGATGTACTTTGTGGAAGATG 1982
Db 2114 GCGACATCAACAAGCTGAAGCCCTGGAGCTCTTTGATGTACTTTGTGGAAGATG 2173
Qy 1983 CTGGCCCATGAGATGCTGCACAGTTTACAGATTTCTGTATCCCGATGCGAATG 2042
Db 2174 CTGGCCCATGAGATGCTGCACAGTTTACAGATTTCTGTATCCCGATGCGAATG 2233
Qy 2043 TCCAGAAAAACGAGCTCAGCTGGCAATGCCCTTCGGCATCTTGGTTCAATTTCT 2097
Db 2234 TCCAGAAAAACGAGCTCAGCTGGCAATGCCCTTCGGCATCTTGGTTCAATTTCT 2288

RESULT 6

US-10-425-114-16424
; Sequence 16424, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E

; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 16424
; LENGTH: 3699
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3063-134-E5_FLI
US-10-425-114-16424

Query Match 96.5%; Score 2023.8; DB 16; Length 3699;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 2028; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 63 TCCGAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCCCTCCACCCGAC 122
Db 199 TTCAGAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCCCTCCACCCGAC 258
Qy 123 ACCACCACCCGACCTTTGCCAGACCCACACCCCGGAGCCAGAGAGGAGATCCT 182
Db 259 ACCACCACCCGACCTTTGCCAGACCCACACCCCGGAGCCAGAGAGGAGATCCT 318
Qy 183 GGGATCAGATGATGAGGACCAAGAGACCTCTCGGACTACTGCAAGGTGGATATCATCC 242
Db 319 GGGATCAGATGATGAGGACCAAGAGACCTCTCGGACTACTGCAAGGTGGATATCATCC 378
Qy 243 AGTGAATTTGAGACCTCTTCAATGGCCGGTATCATTTATTAGAAAGCTTGGATGGGG 302
Db 379 AGTGAATTTGAGACCTCTTCAATGGCCGGTATCATTTATTAGAAAGCTTGGATGGGG 438
Qy 303 GCATTTCTACTGTCTGCTGTGGATATGAGGGGAAAGATTTGTTGCAATGA 362
Db 439 GCATTTCTACTGTCTGCTGTGGATATGAGGGGAAAGATTTGTTGCAATGA 498
Qy 363 AGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAAATAAATTTGCTCAA 422
Db 499 AGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAAATAAATTTGCTCAA 558
Qy 423 ATGTGTTGAGAAAGTATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA 482
Db 559 ATGTGTTGAGAAAGTATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA 618
Qy 483 CGACTTCAAGATTTGAGGATGAATGGGATACATGCTGCTGATGGTCTCGAAGTACTTGG 542
Db 619 CGACTTCAAGATTTGAGGATGAATGGGATACATGCTGCTGATGGTCTCGAAGTACTTGG 678
Qy 543 CCACCATCTCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCTCCCAAGTACTTGTGT 602
Db 679 CCACCATCTCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCTCCCAAGTACTTGTGT 738
Qy 603 GAAGAGTATCATTCGACAGGCTCTTCAAGGGTTAGATTACTTACACAGTAAGTCAAGAT 662
Db 739 GAAGAGTATCATTCGACAGGCTCTTCAAGGGTTAGATTACTTACACAGTAAGTCAAGAT 798
Qy 663 CATTCATCTGACATAAAGCCGGAATAATCTTGTGTTGGATGATGCATATGTGAG 722
Db 799 CATTCATCTGACATAAAGCCGGAATAATCTTGTGTTGGATGATGCATATGTGAG 858
Qy 723 AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGTCTCTCTCTCCCTCCAGGGTC 782
Db 859 AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGTCTCTCTCTCCCTCCAGGGTC 918
Qy 783 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTTAGGAAAAATATCTTAAAAACAAAAAGAA 842
Db 919 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTTAGGAAAAATATCTTAAAAACAAAAAGAA 978
Qy 843 AAAAAGTAAAAAGAAACAGAGAGGAGGCTGAGTTATTGGAGAGGCGCTCGCAGAGAT 902

Db 697 CCACCATCTCTCAAGTGGATCATCAAAATCCAATATCAAGGCCTCCACAGTACGTTGTGT 756
Qy 603 GAAGAGTATCATTCGACAGCTCTTCAAGGGTAGTACTTACACAGTAAGTGCAGAT 662
Db 757 GAAGAGTATCATTCGACAGCTCTTCAAGGGTAGTACTTACACAGTAAGTGCAGAT 816
Qy 663 CATTCATCTGCATATAAGCCGGAATAATCTTGATGTGTGGATGATGCATATGTAG 722
Db 817 CATTCATCTGCATATAAGCCGGAATAATCTTGATGTGTGGATGATGCATATGTAG 876
Qy 723 AAGATGGCAGCTGAGGCCACCTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCAAGGTC 782
Db 877 AAGATGGCAGCTGAGC---CTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTTCAAGGTC 933
Qy 783 TCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTTAAAAACAAAAGAA 842
Db 934 TCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTTAAAAACAAAAGAA 993
Qy 843 AAAAAGTAAAAAGAAACAGAGAGGCGAGCTGAGTTATTGGAGAGCGGCTGCAGAGAT 902
Db 994 AAAAAGTAAAAAGAAACAGAGAGGCGGCTGAGTTATTGGAGAGCGGCTGCAGAGAT 1053
Qy 903 AGAAGATTTGGACGAGAGCTGAAAGGAAAAATATAGAGAAAAACATCACCTCAGCTGC 962
Db 1054 AGAAGATTTGGACGAGAGCTGAAAGGAAAAATATAGAGAAAAACATCACCTCAGCTGC 1113
Qy 963 ACCTTCCATGACCGAGATGGCGAATACCTGCCAGAGGTGAACTTAAAAACAAAGGAT 1022
Db 1114 ACCTTCCATGACCGAGATGGCGAATACCTGCCAGAGGTGAACTTAAAAACAAAGGAT 1173
Qy 1023 AGAGAGGCGGCTGAGGACAGAGCTGCAAGAGCAATGCTGAAGCTGAGGACACAGAA 1082
Db 1174 AGAGAGGCGGCTGAGGACAGAGCTGCAAGAGCAATGCTGAAGCTGAGGACACAGAA 1233
Qy 1083 GAAAGAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGTAGATCAGGAAT 1142
Db 1234 GAAAGAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGTAGATCAGGAAT 1293
Qy 1143 TCGGAACATAGACCTTACCTGATAGATACCTTAAACCAATGGCCATATTTAGAGATGG 1202
Db 1294 TCGGAACATAGACCTTACCTGATAGATACCTTAAACCAATGGCCATATTTAGAGATGG 1353
Qy 1203 CCATTTCTCAGTGGAGCACTGAGCAATGAGATGATGATGAAGAGCTGCCCAA 1262
Db 1354 CCATTTCTCAGTGGAGCACTGAGCAATGAGATGATGATGAAGAGCTGCCCAA 1413
Qy 1263 TCTGAGGAATATAATCTTGTAGTGGCCAAATGAGAAAGTGAATACACATATAGCAGCTC 1322
Db 1414 TCTGAGGAATATAATCTTGTAGTGGCCAAATGAGAAAGTGAATACACATATAGCAGCTC 1473
Qy 1323 CTATGAACAAATCAATGGTGAATTTGCCAAATGGAACGACATATATATATATATATAT 1382
Db 1474 CTATGAACAAATCAATGGTGAATTTGCCAAATGGAACGACATATATATATATATATAT 1533
Qy 1383 CCCAGAGTTTCCACCTCGTGTCTCTGATGCTTGAACCTGTAGAACCTGTGGCTGCGCTCTGT 1442
Db 1534 CCCAGAGTTTCCACCTCGTGTCTCTGATGCTTGAACCTGTAGAACCTGTGGCTGCGCTCTGT 1593
Qy 1443 GCTTTCTGAGGATCACCCTTACTGAGCAAGAGGAGACAGTCCATCCCATGACAGAG 1502
Db 1594 GCTTTCTGAGGATCACCCTTACTGAGCAAGAGGAGACAGTCCATCCCATGACAGAG 1653
Qy 1503 CAGAACGGTTTCAAGCTCAGTACCTGGGATTTGCCAAAGCAAAACCCCGGCGAGCTGA 1562
Db 1654 CAGAACGGTTTCAAGCTCAGTACCTGGGATTTGCCAAAGCAAAACCCCGGCGAGCTGA 1713
Qy 1563 CTTGTTGGTGAATCCCTGGATCCGCGGATGAGATGAGTAAATTTAGATGAAAAATTCCTGA 1622
Db 1714 CTTGTTGGTGAATCCCTGGATCCGCGGATGAGATGAGTAAATTTAGATGAAAAATTCCTGA 1773
Qy 1623 CTTGGGAATGCTTGTGTTGGTGCATAAAACACTTCCGGAAGACATCCAGACGCGTCAGTA 1682
Db 1774 CTTGGGAATGCTTGTGTTGGTGCATAAAACACTTCCGGAAGACATCCAGACGCGTCAGTA 1833

Qy 1683 CCGCTCCATAGAGGTTTAAATAGGAGCGGGTACACAGACCCCTCGGACATCTGGAGCAC 1742
Db 1834 CCGCTCCATAGAGGTTTAAATAGGAGCGGGTACACAGACCCCTCGGACATCTGGAGCAC 1893
Qy 1743 GCGGTGTATGGCATTTGAGCTGGCAACGGGAGATTTATTTGTTGAACCAATTTCTGGGA 1802
Db 1894 GCGGTGTATGGCATTTGAGCTGGCAACGGGAGATTTATTTGTTGAACCAATTTCTGGGA 1953
Qy 1803 AGATATTTCAGAGAGAGAACCAATAGCCCCACATCATAGAGCTGCTAGGAGATATTC 1862
Db 1954 AGATATTTCAGAGAGAGAACCAATAGCCCCACATCATAGAGCTGCTAGGAGATATTC 2013
Qy 1863 AAGGACATTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTCGGAAATTTCTCAATCGCAGAGGAACT 1922
Db 2014 AAGGACATTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTCGGAAATTTCTCAATCGCAGAGGAACT 2073
Qy 1923 GCGACACATCACCAAGCTGAAGCCCTGGAGCCCTTTGATGTACTTGTGAAAAAGTATGG 1982
Db 2074 GCGACACATCACCAAGCTGAAGCCCTGGAGCCCTTTGATGTACTTGTGAAAAAGTATGG 2133
Qy 1983 CTGGCCCCATGAAGATGTGCAAGTTTACAGATTTCTCGATCCCGATGTTAGAAATGT 2042
Db 2134 CTGGCCCCATGAAGATGTGCAAGTTTACAGATTTCTCGATCCCGATGTTAGAAATGT 2193
Qy 2043 TCCAGAAAAACGAGCTCAGCTGGCGAATGCTTGGCATCTCTGTTGAATCT 2097
Db 2194 TCCAGAAAAACGAGCTCAGCTGGCGAATGCTTGGCATCTCTGTTGAATCT 2245

RESULT 8

US-09-880-107-3294
; Sequence 3294, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3294
; LENGTH: 4326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U09564
US-09-880-107-3294

Query Match 30.1%; Score 631.4; DB 9; Length 4326;
Best Local Similarity 60.4%; Pred. No. 7.1e-156;
Matches 1170; Conservative 0; Mismatches 671; Indels 96; Gaps 4;
Qy 160 CCGAGAGCCAGAGGAGAGATCTGGGATCAGATGATGAGGACAGAGACCCCTGGGAC 219
Db 232 CCAGAGCAGGAGAGAGATCTGGGATCTGATGATGATGAGCAAGAGATCTTAATGAT 291
Qy 220 TACTTCAAGGTGGATATCATCCAGTGAATAATTTGAGACCTCTTCAATCGCCGATCAT 279
Db 292 TATTGTAAGGAGGTATCATCTTGTGAAATTTGAGATCTATTCAATGGAGATACCAT 351
Qy 280 GTTATTAGAAAGCTTGGATGGGGCATTCTCTACTGTCTGCTGTCTGCTGGATATGAC 339
Db 352 GTGATCCGAAAGTTAGGCTGGGGACATTTTCAACAGTATGTTATCATGGGATATTCAG 411

QY 340 GGGAAAGATTGTTGCAATGAAGTTGTAAAGAGTCCCGCAGATTATACGAGACAGCC 399
DB 412 GGGAAAGATTGTTGCAATGAAGTTGTAAAGAGTCCCGCAGATTATACGAGACAGCA 471
QY 400 TTGGATGAATTAATTTGCTCAATGTTGTTGAGAAAGTATCCAGTGCAGCCCAACAAA 459
DB 472 CTAGATGAATCCGGTGTGCTGAAGTCAGTTGCGCAATTCAGACCCCTTAATGATCCAAATAGA 531
QY 460 GACATGTTGGTCCAGCTCATTGACGACTTCAAGATTTCAGGCGCATCAATGGGATACATGTC 519
DB 532 GAAATGGTTGTTCAACTAGATGACTTTAAATATCAGGTTTAATGGACACATATC 591
QY 520 TGCATGTTCTTCGAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAATCCAATCAT 579
DB 592 TGCATGTTTGAAGTTTGGGGCATCATCTGCTCAAGTGGATCATCAATCCAATCAT 651
QY 580 CAAAGGCTCCAGTACGTTGTTGAAGTATCATTCGACAGTTCATTCGACAGTTCCTCAAGGTTAGAT 639
DB 652 CAGGGGCTTCCACTGCTTGTGTCAAAATAATTTATTCAGCAAGTGTATACAGGTTCTTAT 711
QY 640 TACTTACACAGTAAGTGCAGATCATCTCATCTGATATAAGCCCGGAAATATCTTTGATG 699
DB 712 TATTTACATACCAAGTGCAGTATCATCTCATCTGATATAACCAAGAGACATCTTTATG 771
QY 700 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 759
DB 772 TCAGTGAATGAGCAGTACATTCGAGGCTGGTGCAGAAAGCAAGAAATGGCAGGCTCT 831
QY 760 GTGTCT 819
DB 832 GGAGCT 891
QY 820 AAAATATCTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 879
DB 892 AAAATGTCAGAAATGAAGAAAGAAATTTGAAGAAAGAAAGAAAGAAAGAAAGAAAG 951
QY 880 TTGGAGAGCGCTCTCAGAGATAGAAGAAATTTGGAGCGAGAAAGCTGAAGAGAAATA 939
DB 952 CTAGAGAGCGAATGAGGAAATTTGAGGAAATTTGAGGAAATTTGAGGAAATTTGAG 997
QY 940 GAAGAAACATCACCTCAGCTGCACCTTCCATGACCAGGATGGCGAATCTGCCAGAG 999
DB 998 -----GCCCTGGGCAAAAAGACCAAAAC 1020
QY 1000 GTGAACTAAAAACAGGATTAGAGAGGGGCTGAGGCGAGAGACTCAAGAGGACAAAT 1059
DB 1021 AAGCAAGAGATCAGAGATCTCTGTTGAAGACCTTGAAGAGAGACCCCACTAATAA 1080
QY 1060 GTGAAGCTGAGGACCCAGAGAGAGAAAGATGCTGAGAAAGAAACATTTGAAAGAT 1119
DB 1081 ATGACCAAGAAAACTTGAAGAGTCAAGTACCATTGGCCAGGATCAAAACGCTTATGGAA 1140
QY 1120 GAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1179
DB 1141 CGTGATACAGAGGTTGGTGCAGAGAAATTAATTTGAATGGAGTGAATGAAGTCAATTA 1200
QY 1180 ACCAATGGCCATATTCAGAAATGGCCCATCTCACTGGAGCAGCAATTCGACGATGAAGAT 1239
DB 1201 TATACTCAGAACAGTAAATGAACATTCAGACATTAAGAGGATCTACATAATGCTAAT 1260
QY 1240 GATGATGAAGAGACTGCCCCAATCTGAGGAATATAATCTTGATGAGGCCCAATGAGAA 1299
DB 1261 GACTGTGATGTCAAAAATTTGAATCAGGAATCTAGTTTCTTAAGTCTCCCAATGGAGA- 1319
QY 1300 AGTGATTACATATAGCAGCTCTTATGAACATTCATGATGATGATGATGATGATGATGAT 1359
DB 1320 -----CAGCAGCAATCTCAAGAAACAGACTCTTGTGACCTATTAACA 1362
QY 1360 CATAAAAATCCCGAGTTCAGAGTTTCCAGAGTTTTCACCTCTGTTGTTCTCTGATCCTTA 1419
DB 1363 TCTGAGGTGTCACACCATGGTGTGCCAGTCTTCTCAACTGTAGTCACTTCACTCACT 1422
QY 1420 GAACCTGTGGCTGTGGCTCTGTGCTTTCTGAGGGATCACCATTACTGAGCAGAGGAG 1479

DB 1423 GAAC-----AACACATTAGCCAACTTCAAGAAAGCATTCGGGC 1460
QY 1480 AGCAGTCCATCCCATGACAGAGCAGAAACGTTTTCAGCTCCAGTACTTGGGGATTTGCCA 1539
DB 1461 AGAGATACCTGTGCAAGATGAACAAGAGCAAGACATAAACGACCACTGGACA-----AC 1515
QY 1540 AAAGCAAAAAACCCGGGAGCTGACTTGTGTTGTTGATCCCTGGATCCCGGAAATGACAGAT 1599
DB 1516 AAAGGAAAAATCCACGGCTGGAAATTTTCTTGTAAATCCCTTGGACCAAAAAATGCGAAA 1575
QY 1600 AAAATTAGATGAATAAATTTGCTGACCTGGGAAATCTTGTGGGTGTCATAACACTTCACG 1659
DB 1576 AAGCTCAAGGTGAAGATTTGCTGACCTTGGAAATCTTGTGGGTGTCACAAACATTTCACT 1635
QY 1660 GAAGACATCCAGACGCGTCAGTACCGCTCCATAGAGGTTTTAAATAGGAGCGGGGTACAGC 1719
DB 1636 GAAGATATCAACAAGCAATATCTTCTTGGAGTTCTAATCGATCTGGCTATAAT 1695
QY 1720 ACCCTGCGGACATCTGGAGCAACGCGGTGATGAGATTTGAGCTGGCAACGGGAGATTAT 1779
DB 1696 ACCCTGCTGACATTTGGAGCACGCGATGCTATGGCTTTGAACTGGCCACAGGTGACTAT 1755
QY 1780 TTGTTTGAACACATTTCTGGGGAAGACTATTTCCAGAGACGAAGACCATAGCCACATC 1839
DB 1756 TTGTTTGAACCTCATTTAGGGGAAGATGACTCGAGATGAAGATCACTTTGCAATTGATC 1815
QY 1840 ATAGAGCTGTAGGAGGATTTTCAAGGCACTTTCTCTATCTGGAATAATTTTCGGGAA 1899
DB 1816 ATAGAACTTCTGGGGAAGGTGCTCGCAAGCTCATTTGGCGAGGAAATAATTCGAAGAA 1875
QY 1900 TTCTTCAATCCAGAGAGAGAACTGGGACACATCAACAGCTGAAGCCCTGGAGCCTCTTT 1959
DB 1876 TTTTTCACAAAAAAGGTGACTGAAACATATACAGAGCTGAAACCTTTGGGGCTTTTT 1935
QY 1960 GATGATCTTGTGGAAGATGAGTGGCTGGCCCATCAAGATGCTGCACACTTTACAGATTTC 2019
DB 1936 GAGGTTCTAGTGGAGAGATGATGAGTGGTCTCAGGAAGAGGAGCTGGCTTTCACAGATTTC 1995
QY 2020 CTGATCCCGATGTTAGAAATGTTTCCAGAAAAACGAGCTCAGCTGGCGAAATGCTTCGG 2079
DB 1996 TTACTGCCCATGTTGGAGCTGATCCCTGAGAAGAGAGCCACTGCCCGGAGTGTCTCCGG 2055
QY 2080 CATCCTTGGTTGAATTC 2096
DB 2056 CACCCTTGGCTTAATC 2072

RESULT 9

US-10-172-118-875
; Sequence 875, Application US/10172118
; Publication No. US20030224374A1

GENERAL INFORMATION:

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene

; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients

; FILE REFERENCE: 9301-175-999

; CURRENT APPLICATION NUMBER: US/10/172,118

; PRIORITY FILING DATE: 2002-06-14

; PRIOR APPLICATION NUMBER: 60/380,770

; NUMBER OF SEQ ID NOS: 2699

; SEQ ID NO 875

; LENGTH: 4326

; TYPE: DNA

; ORGANISM: Homo sapiens

; PUBLICATION INFORMATION:

```

; DATABASE ACCESSION NUMBER: NM_003137
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-875

```

Query Match	30.1%	Score 631.4	DB 15	Length 4326
Best Local Similarity	60.4%	Pos. No. 7.1e-156		
Matches 1170	Conservative	0	Mismatches 156	Indels 96
				Gaps 4
Qy	160	CCGAGAGCCAGAGGAGAGATCCTGGGATCAGATGATGAGGAGCAAGAGAGACCTCGCGAC	219	
Db	232	CCAGAGCAGGAAGAGGAGATCTGGGATCTGATGATGATGAGCAAGAGATCCTAATGAT	291	
Qy	220	TACTGCAAAAGGTGGATATCATCCAGTGAATAATGGAGACCTCTTCAATGGCGGTATCAT	279	
Db	292	TATTGTAAAGAGAGGTTATCATCTTGTGAAAAATGGAGATCTATTCAATGGGAGATACCAT	351	
Qy	280	GTTATTAGAAAGCTTTGGATGGGGGCACCTTCTCTACTGTCTCGCTGTCTGGGATATGCG	339	
Db	352	GTGATCCGAATAGTTAGGCTGGGACACTTTTCAACAGTATGGTTATCATGGGATATTG	411	
Qy	340	GGGAAAAAGATTGTTGCAATGAAAAGTTGTAAAAAGTGCACAGCATTTATACGGAGACAGCC	399	
Db	412	GGGAAGAAATTTGTGGCAATGAAAGTAGTTAAAGAGTCTGAACATTACACTGAAACAGCA	471	
Qy	400	TTGGATGAAATAAAAATTTGCTCAAAATGTGTTGCGAGAAAGTGATCCAGTGACCCAAACAA	459	
Db	472	CTAGATGAAATTCGGGTTTGCTGAAGTCAGTTGCGCAATTTCCAGACCCCTAATGATCCAAATAGA	531	
Qy	460	GACATGGTGGTCCAGCTCATTTGACGACCTTCAAGATTTTCAGGCATGAATGGGATACATGTC	519	
Db	532	GAAATGGTTGTTCAACTACTAGATGACCTTTAAATATACGAGGTTAATGGAACACATATC	591	
Qy	520	TGCATGTTCTTGAAGTACTTTGGCCACCATCTCTCAAGTGGATCATCAAAATCCAACATAT	579	
Db	592	TGCATGTTATTTGAAGTTTTGGGTCATCATCTGCTCAAGTGGATCATCAAAATCCAATTTAT	651	
Qy	580	CAAGGCTCCCACTACGTTGTTGTAAGATGATCATTCGACAGGTCTTCAAGGGTTAGAT	639	
Db	652	CAGGGGCTTCCACTGCGCTTGTGTCAAAAAAATTAATTCAGCAAGTGTGTTACAGGGTCTTGAT	711	
Qy	640	TACTTTACACAGTAAAGTCAAGATCATTTCAATCTGACATAAAGCCGGAAAAATATCTTTGATG	699	
Db	712	TATTTACATACCAAGTCGGTATCATCCACACTGACATTAACACAGAGAACATCTTATTG	771	
Qy	700	TGTTGTGATGATGCATATGTGAGAAAGATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCA	759	
Db	772	TCAGTGAATGAGCAGTACATTTCCGAGGGTGGCTGCAGAAAGCAACAGAAATGGCAGCGATCT	831	
Qy	760	GGTCTCTCTCTCTTCAAGGCTGTCAGTGTAGTACGGCTCCACAGCAGAGAACTATAGGA	819	
Db	832	GGAGCTCTCTCGGCTTTCCGATCTGCACTGAGTACTGTCTCCCAAGCTTAACAGCTGAC	891	
Qy	820	AAAAATATCTAAAAACAAAAAGAAAAATCTAAAAAGAAACAGAAAGCGCAGGCTGAGTTA	879	
Db	892	AAAAATGTCAAGAATAAGAAAGAAATTTGAAGAAGAAAGCAGAAAGCGCCAGCAGAAATTA	951	
Qy	880	TTGAGAAAGCCCTGCAGGAGATAGAGAAATTTGAGCGGAGAAAGCTGAAGGAAATTAATA	939	
Db	952	CTAGAGAAAGCGAATTCAGGAAATTTGAGGAAATTTGAGGAAATGGGAAAGAGTCCG	997	
Qy	940	GAAGAAAAACATCACCTCAGCTGCACCTTTCCAAATGACACAGGATGCGGAATACTGCCCAGAG	999	
Db	998	-----GCCCTGGGCAAAAAAGACCAAC	1020	
Qy	1000	GTGAAACTTAAAAACAACAGGANTTAGAGAGCGGCTGAGGCAGAGACTGCAAGAGGACAAT	1059	
Db	1021	AAGCAAGAAAGAAATCAGAGAGTCTGTTGAAAGAACCCTTGAAGAGAGAACCCACCTAAATAA	1080	
Qy	1060	GGTGAAGCTGAGGACCGAGGAGAGAAAGAAAGATGCTGAGAAAGAAACATTTGAANAAGAT	1119	
Db	1081	ATGACCCAAGAAAACTTTGAAGATTCAGTATACCATTTGGCCAGGATCAAAACGCTTATGGA	1140	
Qy	1120	GAAGATGATGTAGATCAGGAATTTGCGAAACATAGACACCTCAGTGGATAGATACCTAAA	1179	

RESULT 10
US-10-342-887-875
; Sequence 875, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:

```

; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 875
; LENGTH: 4326
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-342-887-875

Query Match      30.1%; Score 631.4; DB 16; Length 4326;
Best Local Similarity 60.4%; Pred. No. 7.1e-156; Indels 96; Gaps 4;
Matches 1170; Conservative 0; Mismatches 671;

QY 160 CCGAGCCAGAGAGAGATCTGGGATCAGATGATGAGGACCAAGAGACCTCCGGAC 219
DB 232 CCAGACAGGAGAGAGATCTGGGATCTGATGATGATGAGGAGATCCTAATGAT 291
QY 220 TACTGCAAGGTGGATATCATCCAGTGAATTTGGAGCTCTCAATGGCCGGTATCAT 279
DB 292 TATTGTAAGGAGGTATCATCTGTGAAATTTGGAGATCTATCAATGGGAGATACCAT 351
QY 280 GTTATTAGAAAGCTTCGATGGGGCACCTCTACTGTCTGGCTGTGGGATATGACG 339
DB 352 GTGATCCGAAAGTTAGCTGGGACACTTTCAACAGTATGTTATCATGGGATATTCAG 411
QY 340 GGGAAAGATTTCTGCAATGAAAGTTGAAAGTCCAGGATATATACGAGACAGCC 399
DB 412 GCGAAGAAATTTGTGCAATGAAAGTAGTTAAAGTGTGAAATTTACACTGAAACAGCA 471
QY 400 TTGGATGAATATAAATGTCTCAATGTGTTGAGAAAGTATCCAGTGCACCCAAACAA 459
DB 472 CTAGATGAATACCGGTTGCTGAAAGTCAGTTTCGCAATTCAGACCTAATGATCCAAATAGA 531
QY 460 GACATGGTGTCCAGCTCATTTGACGACTTCAAGATTTTCAAGGATGATGGGATACATGTC 519
DB 532 GAAATGGTTGTTCAACTACTAGATGACTTTTAAATATATCAGGAGTTAATGGAAACACATATC 591
QY 520 TGCAATGTCTTCAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAATCCAACTAT 579
DB 592 TGCATGTTATTTGAAGTTTGGGGCATCATCTGCTCAAGTGGATCATCAATCCAAATAT 651
QY 580 CAGGCTCTCCAGTACGTTGTGAGAGATCATTCGACAGTCTCTTCAAGGTTTAGAT 639
DB 652 CAGGGCTTCCATGCTGTTGTCAAAAAATTTATTCAGCAAGTTTACAGGCTCTTGAT 711
QY 640 TACTTACACAGTAAAGTCAAGATCATTTATCATGACATAAAGCCGGAATAATATCTTGATG 699
DB 712 TATTTACATACCAAGTCCGATCATCCACACTGACATTAATTAACCCAGAGAAACATCTTATTG 771
QY 700 TGTGTGATGATGATATGTCAGAGAAATGGCAGCTGAGCCACTGAGTGGCAGAAAGCA 759
DB 772 TCAGTGAATGAGCATATTCGGAGGCTGCTGTCGAGAGCAACAGAAATGGCAGCATCT 831
QY 760 GGTGCTCTCTCTCTTCAAGGCTCTGAGTGAAGTACGGCTCCACAGCAGAAACCTATAGGA 819
DB 832 GGAGCTCTCTCGCTTTCGGATCTGAGTCACTGCTCTCCAGCTTAAACCCAGCTGAC 891

```

```

QY 820 AAAATATCTAAAAACAAAAAGAAAAAACTGAAAAAGAAAAACAGAGAGGCGCTGAGTTA 879
DB 892 AAAATCTCAAAGAAATTAAGAAAGAAATTAAGAAAGAAAGCAGAGAGCGCCAGGAGAAATTA 951
QY 880 TTGGAGAGCGCTCGCAGGAGATAGAAATTTGGAGCGAGAGAGCTGAAAGGAAAAATAAATA 939
DB 952 CTAGAGAGCGAAATGCGAGAAATTTGAGGAAATGAGAAAGAGTCGG----- 997
QY 940 GAAGAAAAACATCACCTCAGCTGCACCTTCCAATGACCCAGGATGGCGAATATCTGCCAGAG 999
DB 998 -----GCCCTGGGCAAAAAAGACCAAAAC 1020
QY 1000 GTGAAACTAAAAACAACAGGATTTAGAGGAGGCGGCTGAGGCGAGAGACTCCAAAGGACAAT 1059
DB 1021 AAGCAAGAGATCAGAGAGTCTCTGTGAAAGACCTTTGAAAGAGAACCCCACTAATAA 1080
QY 1060 GGTGAAGCTGAGGACGAGGAGAAAGAGATGCTGAGAAAGAAACATTTGAAAAAGAT 1119
DB 1081 ATGACCCCAAGAAAAACTTTGAAAGTCAGTACCATTTGGCCAGGATCAACGCTTATGGAA 1140
QY 1120 GAAGATGATGTAGATCAGGAACTTTGCAACATAGACCTTACGTGGATAGATCACCTAAA 1179
DB 1141 CGTGATACAGAGGTGTGCGAGCAAAATTAATTGCAATGGAGTGAATGAAGTCATTAAT 1200
QY 1180 ACCAATGGCCATTTGGAATGSCCAATTTCTCACTGGAGCAGCAACTTGGACGATGAAGAT 1239
DB 1201 TATACTCAGAACAGTAATAATGAACATTTGAGACATAAAGAGGATCTACATAATGCTAAT 1260
QY 1240 GATGATCAGAAAGACTGCCCAATCTGTAGGAAATATAATCTTGATGAGCCAAATGCAAGAA 1299
DB 1261 GACTGTGATGTCCAAATTTGAAATCAGGAATCTAGTTTCTTAAGTCTCCCAATGGAGA- 1319
QY 1300 AGTGATTTACATATAGCAGCTCCTATGAAACAATTCATGGTGAATTTGCCAATGACGA 1359
DB 1320 -----CAGCAGCACATCTCAAGAAACAGACTCTTGTAACCTTATAACA 1362
QY 1360 CATAAATTTCCGAGTCACAGTTCCAGAGTTTCCACCTCGTTGTTCTCTGTGATCCTTA 1419
DB 1363 TCTGAGGTGTCAGACACCATGTTGTCAGTCTTCTCACTGATGATGATGATGATGATGAT 1422
QY 1420 GAACTGTGGCTCGGCTCTGTCTTTCTGAGGGATCAACCTTACTGAGCAAGAGAG 1479
DB 1423 GAAC-----AACACATTAGCCAACTTCAAGAAAGACTTCGGGC 1460
QY 1480 AGCAGTCCATCCCATGACAGAGCAGACCGTTTTCAGCTCCAGTCTGGGGATTTGGCA 1539
DB 1461 AGAGATACCTGTGAAGATGAACAAGAGCAAGAACATAACGACCACTTGGACA-----AC 1515
QY 1540 AAAGCAAAAAACCGGCGAGCTGACTTGTGTTGTAATCCCTTGGATCCGCGGAATGCGAGAT 1599
DB 1516 AAAGGAAATCCACGGCTGGAAATTTCTTGTAAATCCCTTTGAGCCAAAAAATGCGAA 1575
QY 1600 AAAATTTAGAGTAAAAATTTGCTGACTGGGAAATGCTGTTGGGTGTCATAACACTTCACG 1659
DB 1576 AAGCTCAAGGTGAAGATTTGCTGACCTTGGAAATGCTTGTGGGTGTCACAAACATTTTCACT 1635
QY 1660 GAAGACATCCAGACCGCTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGGTACAGC 1719
DB 1636 GAAGATATTTCAAAACAAAGGCAATATCGTTCTTGGAAAGTTCTAATCGGATCTGGCTAAT 1695
QY 1720 ACCCTCTGGGACATCTGGAGCAGCGCTGTATGGCAATTTGAGCTGGCAACCGGAGATTAAT 1779
DB 1696 ACCCTCTGTGACATTTGGAGCAGCGCATGCTGCTTGTGAACCTGGCCACAGGTGACTAT 1755
QY 1780 TTGTTTGAACCACTTTCTGGGGAAGACTATTTCCAGAGCAGAGACCATAGCCCAACATC 1839
DB 1756 TTGTTTGAACCTCATTTCAAGGGAAGAGTACACTCGAGATGAAGATCACTATGCTATGATC 1815
QY 1840 ATAGAGCTCTAGGAGTATTTCCAAAGGCACTTTGCTCTATCTGGAATAATATTTCTCGGAA 1899
DB 1816 ATAGAACTTTCTGGGAGAGTGTCTCGAAGCTCATTTGTGGCAGGAAATATTTCCAAGGAA 1875
QY 1900 TTCTTCAATCGCAGAGGAGAACTGCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTT 1959

```

Db 1876 TTTTCCACCAAAAGGTGACCTGAAACATATACAGAGCTGAAACCTTGGGCGCTTTT 1935
Qy 1960 GATGTAATTTGTGAAAGATATGCTGGCCCATGAAGATGCTGCACAGTTTACAGATTC 2019
Db 1936 GAGGTTCTAGTGAGAGATGATGAGTGTCTCAGGAAGAGGACGCTGGCTTACAGATTC 1995
Qy 2020 CTGATCCCGATGTTAGAAATGGTTCCAGAAAAACGAGCCCTCAGCTGGCGAATCCCTTCGG 2079
Db 1996 TTACTGCCCATGTTGAGCTGATCCCTGAGAGAGAGCCACTGCCCGGAGTGTCTCCGG 2055
Qy 2080 CATCTTGGTTGAATTC 2096
Db 2056 CACCCTTGGCTTAATC 2072
RESULT 11
US-10-305-720-703
; Sequence 703, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
; SEQ ID NO 703
; LENGTH: 923
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 307624
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1) ... (923)
; OTHER INFORMATION: a, t, c, g, or other
US-10-305-720-703
Query Match 30.0%; Score 629.4; DB 16; Length 923;
Best Local Similarity 88.3%; Pred. No. 1.1e-155;
Matches 700; Conservative 0; Mismatches 89; Indels 4; Gaps 3;
Qy 67 AAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCCAGCGCCACCA 126
Db 43 AANAAGGCGCCCTGGCCAAAGCTCANTTTTCTCTCTCCAGCGCCACCA 102
Qy 127 CCACCAACCGCCACCTTTGGCAGACCCCAACCCCGGAGCCAGAGGAGATCTCTGGGA 186
Db 103 CCACCAACCGCCCTTTGGCAGACCCCAACCCCGGAGCCAGAGGAGATCTCTGGGA 162
Qy 187 TCAGATGATGAGGAGCAAGAGCCCTGCGGACTCTCAAGGTGATATCATCAGTG 246
Db 163 TCAGATGATGAGGAGCAAGAGCCCTGCGGACTCTCAAGGTGATATCATCAGTG 222
Qy 247 AAAATTGGAGACCTCTTCAATGCGCGTATCATGTTATTAGAAAGCTTGGATGGGGGCAC 306
Db 223 AAAATTGGAGACCTCTTCAATGCGCGTATCATGTTATTAGAAAGCTTGGATGGGGGCAC 282
Qy 307 TTCTCTACTGTCTGGCTGTCTGGGATATGCGAGGGGAAAGATTTGTTCAATGAAGTT 366
Db 283 TTCTCTACTGTCTGGCTGTCTGGGATATGCGAGGGGAAAGATTTGTTCAATGAAGTT 342
Qy 367 GTAAAAAGTCCCGACGATTTATCGGAGACAGCCTTGGATGAATAAATAATGCTCAAAATG 426
Db 343 GTAAAAAGTCCCGACGATTTTNTGTATACAGCCTTGGATGAATAAATAATGCTCAAAATG 402
Qy 427 GTTCGAGAAAGTATCCAGTGCACCAACAGACATGGTGTCCAGCTCATTCAGAC 486

Db 403 GTTCGAGAAAGTATCCAGTGACCCCAACAAAGACATGGTGTCCAGCTCATTCAGCAC 462
Qy 487 TTCAAGATTTTCAGGCAATGAATGGGATACATGTCTGATGGTCTTTCGAAGTACTTGGCCAC 546
Db 463 TTTAAGATTTTCAGGCAATGAATGGGATACATGTCTGATGGTCTTTCGAAGTACTTGGCCAC 522
Qy 547 CATCTCCTCAAGTGGATCATCAAACTCAACTATCAAGGCTCCAGTACGTTGTGTGAAG 606
Db 523 CATCTCCTCAAGTGGATCATCAAACTCAACTATCAAGGCTCCAGTACGTTGTGTGAAG 582
Qy 607 AGTATCATTCGACAGGCTCTTCAAGGTTAGATTTACTTACACAGTAAG-TGCAAGATCAT 665
Db 583 AGTATCATTCGACAGGCTCTTCAAGGTTAGATTTACTTACACAGTAAGTTGCAAGATCAT 642
Qy 666 TCATACTGACATAAAGCCGAAAAATATCTTGTGTGTGGATGATGATATGTGAAG 725
Db 643 ACNTACTGACATAAAGCCGAAAAATATCTTGTGTGTGTGTAAGANGCATATTTAANAAG 702
Qy 726 AATGGCAGCTGAGGCCACTGAGTGGCAGAAACAGGTGCTCTCTCTCTCAGGGTCTGC 785
Db 703 AATGGCAGCTGAG--CCCTGATTTGGCANAAGCAGGTGCTCC-NCTCCCTCAGGGTCTGC 759
Qy 786 AGTGAATGAGGCTCCACAGCAGAAACCTATAGGAAAAATATCTTAAACACAAAAAGAAAA 845
Db 760 ANTGAATTACGCTCCCGCNCNAAACTNTNGGAAAAATTTCTAAACCCNNNGAAAAACTGA 819
Qy 846 ACTGAAAAAGAAA 858
Db 820 AANAACCNAAAA 832
RESULT 12
US-10-252-157-374
; Sequence 374, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 374
; LENGTH: 4349
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 474724.5
US-10-252-157-374
Query Match 30.0%; Score 628.2; DB 15; Length 4349;
Best Local Similarity 60.3%; Pred. No. 5e-155;
Matches 1168; Conservative 0; Mismatches 673; Indels 96; Gaps 4;
Qy 160 CCGGAGCCAGAGGAGAGATCCTGGGATCAGATGATGAGGAGCAAGAGGACCTCGCGGAC 219
Db 232 CCAGAGCAGAGAGAGAGATCCTGGGATCTGATGATGAGGAGCAAGAGATCTTAATGAT 291
Qy 220 TACTGCAAGGTGGATATCATCCAGTGAAAAATTTGGAGACCTCTTCAATGGCCGGTATCAT 279
Db 292 TATTGTAAGAGGATTATCATCTTGTGAAAAATTTGGAGATCTATTCAATGGGAGATACCAT 351
Qy 280 GTTATTAGAAAGCTTGGATGGGGGACATTTCTTACTGTCTGGCTGTCTGGGATATGACAG 339
Db 352 GTGATCCGAAAGTTAGGCTGGGGACACATTTTCAACAGATATGGTTATCATGGGATATTCAG 411
Qy 340 GGGAAAAAGTTTGTTCATGAAAGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCC 399

Db	412	GGGAGAAATTTGTGGCAATGAAAGTAGTTAAAGATGCTGAAACATTACACTGAAACAGCA	471
Qy	400	TTGGATGAAATAAAAATTTGCTCAAAATGTGTTTCGAGAAAGTGATCCGAGTCACCCAAACAAA	459
Db	472	CTAGATGAATCCGGTTGCTGAAGTCAGTTTCGCAATTCAGACCCCTAATGATCCAAATAGA	531
Qy	460	GACATGTGTGCAGCTCATGTGACGACTTCGAAGATTTTCAGGCATGAAATGGGATACATGTC	519
Db	532	GAAATGGTTGTTCAACTACTAGATGACTTTTAAATATCAGGAGTTAATGGGAAACACATATC	591
Qy	520	TGCATGTCTTCGAAGTACTTGGCCACCATCTCTCGAGTGGATCATCAAAATCCAACTAT	579
Db	592	TGCATGTGATTTGAAGTTTGGGGCATCATCTGCTCAAGTGGATCATCAAAATCCAATTTAT	651
Qy	580	CAAGGCCTCCAGTAGCTGTTGTGGAAGAGTATCATTCGACAGGTCCCTTCAAGGGTTAGAT	639
Db	652	CAGGGCTTCCACTGCCCTTGTGTCAAAAAATTTATTTCAGCAAGTGTTCACAGGCTCTTGAT	711
Qy	640	TACTTACACAGTAAGTGCACAGATCATTTATCATGTACATAAAGCGGAAAAATATCTTTGATG	699
Db	712	TATTTACATACCAAGTCCGTATCATCCACACTGACATTTAAACAGAGAACATCTTTATTG	771
Qy	700	TGTGTGATGATGATATGTGAGNAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCA	759
Db	772	TCAGTGAATGAGCAGTAGTACATTCGGAGGCTCGCTGCAGAAGCAACAGAAATGGCAGCGATCT	831
Qy	760	GGTGCTCCTCTCTCCTCAGGCTGTCAGTACAGTACGGCTCCACAGCAGAAACCTATAGGA	819
Db	832	GGAGTCTCTCGCCTTCGGATCTGCACTGAGTACTGCTCCCAAGCTTAAACAGCTGAC	891
Qy	820	AAAAATATCTAAAAA CAAAAAGAAAAA CTGAAAAAGAAACAGAAAGGCGAGGCTGAGTTA	879
Db	892	AAAATGTCAAGAAATAAGAAAGAAATTTGAAGAAGAAGCAGAAGCGCCAGGCAGAAATTA	951
Qy	880	TTGAGAAAGCGCCTGCAGGAGATAGAAAGATTTGGAGCGGAGAGCTGAAAGGAAAAATTA	939
Db	952	CTAGAGAAGCGAATGCGAGAAATTTGAGGAAATGGAGAAATGGAGAAAGAGTCGG	997
Qy	940	GAAGAAACATCACCTCAGCTGCACCTTCCAATGACCCAGGATGGCGAATACTGCCAGAG	999
Db	998	-----GCCCTGGCMAAAAGACCAAC	1020
Qy	1000	GTGAAACTAAAAACAACAGGATTTAGAGGAGCGCGCTGAGGCAGAGACTGCAAAAGGACAAT	1059
Db	1021	AAGCAAGAAGATCAGAGAGTCTCTGTTGAAAGCCCTTTGAAGAGAGACCCACCTAATAAA	1080
Qy	1060	GGTGAAGCTGAGGACAGGAAGAGAAAGATGCTGAGAAAGAAATCATTTGAAAGAGAT	1119
Db	1081	ATGACCCCAAGAAAACTTGAAGAGTCAAGTACCATTGGCCAGGATCAACCGCTTATGGAA	1140
Qy	1120	GAAGATGATGTAGATCAGGAACTTGCAGACATAGACCCCTACGTGGATAGAAATCACCTAAA	1179
Db	1141	CGTGATACAGAGGTGTGTCAGCAGAAATTAATTGCAATGGAGTGAATTGAAGTCATTAAAT	1200
Qy	1180	ACCAATGGCCATATTGAGAAATGGGCCATTCTCACTGGAGCAGCAACCTGGGACGATGAAGAT	1239
Db	1201	TATACTCAGACACTAATATGAACATTTGAGACATTAAGAGGATCTACATATATGCTAAT	1260
Qy	1240	GATGATGAAGAAAGACTGCCAAATCTCTGAGGAATATAATCTTGATGAGGCCAAATTCAGAA	1299
Db	1261	GACTGTGATGTCCAAAAATTTGAATCAGGAATCTAGTTTCTTAAGCTCCCAAAATGGAGA	1319
Qy	1300	AGTGATTAACATATAGAGCTCCTATGACAAATTTCAATGGTGAAATTTGCCAAATGGACGA	1359
Db	1320	-----CAGCAGACATCTCAGAAACAGACTCTTGTCACCTATTAAACA	1362
Qy	1360	CATAAAATTCGCCAGTACAGTCCACAGATTTTCCACCTCGTTGTTCTCTGTGATCCCTTA	1419
Db	1363	TCTGAGGTGTGAGACCAATGGTGTGCCAGTCTTCTCTCACTGTAGGTGAGTCACTCAGT	1422
Qy	1420	GAACTGTGGCCTCGGCTCTGTCTTTCTGAGGGATCAACAATTACTGAGCAAGAGAG	1479
Db	1423	GAAC-----AACCATTTAGCCAACTTCAAGAAAGCATTTCCGGC	1460

Qy	1480	AGCAGTCCATCCCATGACAGAACGAGAA	CGGTTTCAGCCTCCAGTACTGGGATTTGCCA	1539	
Db	1461	AGAGATA	CCCTGTGAAGATGAA	CAAGAGCAAGAACATAACGGACCACTGGACA-----AC	1515
Qy	1540	AAAGCAAAAACCCGGGCAGCTGACTTGTGGTGC	AATCCCTCGATCCGCGGAATCAGAT	1599	
Db	1516	AAAGGAAATCCACGGCTGGAAATTTTCTTGT	TAATCCCTTGAGCCAAAAATGCAGAA	1575	
Qy	1600	AAAATTTAGAGTAAAAATTTGCTGACCTTGGG	AAATGCTTGTGGTGCATAAACA	1659	
Db	1576	AAGCTCAAGGTGAAGATTGCTGACCTTGGAA	ATGCTTGTGGTGCACAAACATTTCACT	1635	
Qy	1660	GAAGACATCCAGACGGTCAGTACCGTCCAT	AGAGGTTTTAATAGAGCGGGTACAGC	1719	
Db	1636	GAAGATATTCAAA	CAAGGCAATATCGTTCTTGGAA	1695	
Qy	1720	ACCCCTCGGACATCTGGACACGGCGTGTAT	GGCAATTTGAGCTGGCAACGGGAGATTAT	1779	
Db	1696	ACCCCTGCTGACATTTGGAGCAGGCATG	CGCTTTGAACTGGCCACAGGTGACTAT	1755	
Qy	1780	TTGTTTGAA	CCACATTTCTGGGGAAGACTATTC	CAGAGACGAGACACATAGCCACATC	1839
Db	1756	TTGTTTGAA	CCTCATTCAGGGGAAGAGTACAT	CTCGAGATGAAGATCACATTTGCATTTGATC	1815
Qy	1840	ATAGAGCTGTGAGCAGTATTTCCA	AGGCACCTTGTCTCTATCTCGAAAATATTTCTCGGGAA	1899	
Db	1816	ATAGAACTTCTGGGGAAGTGCCTCG	CAAGCTCATTTGTGGCAGGAAATATTTCCAAGGAA	1875	
Qy	1900	TTCTTCAATCGCAGAGGAGAACTCG	CACACATCACCAAGCTGAAGCCCTGGAGCCTCTTT	1959	
Db	1876	TTTTTACCAAAAAGGTGACCTGAA	ACAATATCACGAAGCTGAAACCTTTGGGGCCTTTT	1935	
Qy	1960	GATGTACTTGTGGAAAAGTATGCTGG	CCCCCATGAAGATGCTGCA	CAGTTTACAGATTTC	2019
Db	1936	GAGGTTCTAGTGAGAAGTATGATGGT	CCAGGAAAGAGGCAGCTGGCTTT	CAAGATTTC	1995
Qy	2020	CTGATCCCGATGTTAGAAAATGGTT	TCAGAAAAACGAGCCCTCAGCTGCGGAAATGCCTTCGG	2079	
Db	1996	TTACTGCCCATGTTGGAGCTGATCC	CTGAGAGAGGCCACTGCCCGGAGTGTCTCCGG	2055	
Qy	2080	CATCCTTGGTTGAATTC	2096		
Db	2056	CACCTTGGCTTAATC	2072		

RESULT 13

RES001 13
US-09-814-353-21987
; Sequence 21987, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 21987

; LENGTH: 4639

; TYPE: DNA

; ORGANISM: Homo sapiens

; NAME/KEY: misc feature

; LOCATION: 4427, 4428, 4429, 4430, 4431, 4432, 4433, 4434, 4435, 4436,

; LOCATION: 4437, 4438, 4439, 4440, 4441, 4442, 4443, 4444, 4445, 4446,

; LOCATION: 4447, 4448, 4449, 4450, 4451, 4452, 4453, 4454, 4455, 4456,

; LOCATION: 4457

; OTHER INFORMATION: n = A,T,C or G

US-09-814-353-21987

Query Match

Best Local Similarity 30.0%; Score 628.2; DB 10; Length 4639;

Matches 1168; Conservative 0; Mismatches 673; Indels 96; Gaps 4;

```
QY 160 CCGAGCCAGAGGAGGATCCTGGATCAGATGATGAGGAGCAAGAGCACCTGGGAC 219
DB 220 CCAGAGCAGAGAGAGATCTGGGATCTGATGATGATGAGGAGATCCTATGAT 279
QY 220 TACTGCAAGGTGGATATCATCCAGTGAAATTTGGAGACCTCTTCAATGGCCGGTATCAT 279
DB 280 TATTGTAAGGAGGTTATCATCTTGTGAAATTTGGAGATCTATTCAATGGGAGATACCAT 339
QY 280 GTTATTAGAAAGTTGGATGGGGGCACTTCTCTACTGTCTGTGCTGGTGTGGGATATGCGAG 339
DB 340 GTGATCCGAAAGTTAGGCTGGGGGACATTTTCAACAGATGATGTTATCATGGGATATTCAG 399
QY 340 GGGAAAGATTTGTCGAATGAAAGTTGTAAGAGTGGCCAGCATATATACGGAGCAGCC 399
DB 400 GGGAAAGATTTGTCGAATGAAAGTTGTAAGAGTGGCCAGCATATATACGGAGCAGCA 459
QY 400 TTGGATGAAATTAATTTGCTCAATGTTTCGAGAAAGTGATCCCACTGACCCCAACAAA 459
DB 460 CTAGATGAAATCCGGTGTGAGTCACTTGCATTCAGACCTTCAATGATCCAAATAGA 519
QY 460 GACATGGTGGTCAAGTCAATGACGATTCGAGATTCAGGATGATGATGATGATGATGATG 519
DB 520 GAAATGGTGTTCACACTAGATGATGATGATGATGATGATGATGATGATGATGATGATG 579
QY 520 TGCATGGTCTTCCAGTACTTGGCCACCATCTCTCAAGTGGATCATCAATCCCACTAT 579
DB 580 TGCATGGTATTTGAAAGTTTGGGGCATCATCTGCTCAAGTGGATCATCAATCCCACTAT 639
QY 580 CAAGGCTCCAGTACGTTGTGAGAGATGATCATTCGACAGTCTCTTCAAGGGTTAGAT 639
DB 640 CAGGGCTTCCAGTCTGCTGTGTCAGAAATTTTCAAGAGTGTTCAGGGTCTTGAT 699
QY 640 TACTTACAGTAAGTGCAGATCATTTCACTGATCAATGATGATGATGATGATGATGATG 699
DB 700 TATTTACATACCAAGTCCGATCATCCACATGACATTAACACAGAGAACATCTTTATG 759
QY 700 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 759
DB 760 TCAGTGAATGAGCAGTACATTCGAGGCTGGCTGCAGAGAACACAGAAATGGCAGCATCT 819
QY 760 GTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 819
DB 820 GGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 879
QY 820 AAAAATCTAAAAAAGAAAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 879
DB 880 AAAATGTCAGAAATTAAGAAAGAAATTTGAAGAAAGAAAGTAAAGTAAAGTAAAGTAAAG 939
QY 880 TTGAGAGAGCGCTGAGGAGATGAGAAATTTGAGAGGAGAGTAAAGTAAAGTAAAGTAAAG 939
DB 940 CTAGAGAGCGAATGAGGAAATTTGAGAAATTTGAGAAATTTGAGAAATTTGAGAAATTTGAG 985
QY 940 GAAGAAAAATCATCCTCAGCTGACCTTCCAATGACCAAGATGGCGAATACTGCCAGAG 999
DB 986 -----GCCCTGGGCAAAAAAGACCAAAAC 1008
```

```
QY 1000 GTGAACTAAACCAACAGGATTAGAGGAGGGCTGAGGAGAGAGACTGCAAGAGCAAT 1059
DB 1009 AAGCAAGAGAAATCAGAGAGTCTGTGTAAGAGACCTTTGAAAGAGAACCCACTATATAA 1068
QY 1060 GGTGAAGCTGAGGACAGGAGAGAGAAAGATGCTGAGAAAGAGAAAAACATTTGAAAAAGAT 1119
DB 1069 ATGACCCCAAGAAAAAATTTGAAGAGTCAGTACCAATTTGGCCAGGATCAAAACGCTTATGGAA 1128
QY 1120 GAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1179
DB 1129 CGTGATACAGAGGGTGGTGCAGCAAAATTTAATTCATTCGAGTGAATGATTCAGATCAAT 1188
QY 1180 ACCAATGGCCATATTGAGAAATGGCCCAATTTCTCAGTGGAGCAGCACTGAGCAGATGAAGAT 1239
DB 1189 TATATCTCAGAACAGTAATAATAAGAAACATTGAGACATAAAGAGGATCTACATATGCTAAT 1248
QY 1240 GATGATGAGAGAGATGAGCCCAATCTGAGGAATATAATCTTGTGAGGAGCAAAATGAGAA 1299
DB 1249 GACTGTGATGTCAAAATTTGATCAGGAATCTAGTTTCTTAAGCTCCCAAAATGGAGA- 1307
QY 1300 AGTGATTAACATATAGCAGCTCTATGAAACAAATTCATGATGATGATGATGATGATGATG 1359
DB 1308 -----CAGCAGCAGATCTCAAGAAACAGACTCTTGTAACACTATAACA 1350
QY 1360 CATAAAAATCCGAGTCCAGAGTTCCAGAGTTTCCACCTCGTTGTTCTCTGGATCCTTA 1419
DB 1351 TCTGAGGTGTGACAGACCATGGTGTGCCAGTCTTCTCAACTGTAGGTGAGTCAATTCAGT 1410
QY 1420 GAACTGTGGCTCGGGCTCTGTCTTCTGAGGAGTACCCTTACTGAGCAAGAGGAG 1479
DB 1411 GAAC-----AACACTAGCCAACTTCAAGAAAGATTCGGGC 1448
QY 1480 AGCAGTCCATCCCATGACAGAAAGCAGACGGTTTTCAGCTCCAGTACTGGGATTTGCCA 1539
DB 1449 AGAGATACCTGTGAAGATGAACAAGAGCAAGAAATACACGACCACTGGACA-----AC 1503
QY 1540 AAAGCAAAAAACCGGGCAGTCACTTGTGATGATGATGATGATGATGATGATGATGATGATG 1599
DB 1504 AAAGGAAATTCACGGCTGGAAATTTTCTGTGTAATCCCTTTGAGCAGCAAAATGAGAA 1563
QY 1600 AAAATTTAGAGTAAATTTGCTGACCTGGGAAATGCTTGTGGTGGTGCATAAACACTTCAG 1659
DB 1564 AAGCTCAAGGTGAGATGCTGACCTTGGAAATGCTTGTGGTGGTGCATAAACACTTCAT 1623
QY 1660 GAAGACATCCAGACCGCTCAGTACCGCTCCATAGAGGTTTAAATAGGAGCGGGTACAGC 1719
DB 1624 GAAGATATCAACAAAGCAATATCGTTCCTTGGAAATTTCTAATCGGATCTGGCTATAAT 1683
QY 1720 ACCCTCGGACATCTGGAGCAGCGGTGATGATGATGATGATGATGATGATGATGATGATGAT 1779
DB 1684 ACCCTCGTACATTTGGAGCAGCGCATGATGATGATGATGATGATGATGATGATGATGATG 1743
QY 1780 TTGTTTGAACCACTTCTGGGAGAGTATTTCCAGAGAGAGAGTATTTCCAGAGAGAGAGAG 1839
DB 1744 TTGTTTGAACCTTCTAGGGGAGAGTATTTCCAGAGAGAGAGTATTTCCAGAGAGAGAG 1803
QY 1840 ATAGAGCTGCTAGGAGTATTTCCAAAGGCACTTTGTCTCTATCTGGAATATTTCTCGGAA 1899
DB 1804 ATAGAACTTCTGGGAGAGTGTCTGCAAGCTCATTTGTGGCAGGAAATATTTCCAAGGAA 1863
QY 1900 TTCTTCAATTCGAGAGGAGAACTGGGACACATCAACCAAGCTGAAAGCTTGGAGCTCTTT 1959
DB 1864 TTTTTCACCAAAAAAGTGACCTGAAACATATCAAGAGCTGAAACCTTTGGGGCTTTT 1923
QY 1960 GATGACTTCTGGAAGATGATGGCTGGCCCATGATGATGATGATGATGATGATGATGATGAT 2019
DB 1924 GAGGTTCTAGTGAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1983
QY 2020 CTGATCCCGATCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2079
DB 1984 TTACTGCCCATGTTGAGGCTGATTCCTTGAGAAAGAGGCACTTGCCTGCCGAGTGTCTCCGG 2043
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2005, 05:33:48 ; Search time 164.232 Seconds
(without alignments)
9075.705 Million cell updates/sec

Title: US-10-799-676-1_COPY_91_2187
Perfect score: 2097
Sequence: 1 atgagctcccgaaagtgtct.....ggcatccttggtgaattct 2097

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2097	100.0	3253	4	US-09-759-359A-1
2	2097	100.0	3253	4	US-10-207-973-1
3	2023.8	96.5	2791	2	US-09-016-000-8
4	631.4	30.1	4299	1	US-08-264-002-1
5	629.4	30.0	923	4	US-09-016-434-703
6	486.8	23.2	90541	4	US-09-759-359A-3
7	486.8	23.2	90541	4	US-10-207-973-3
8	291.4	13.9	1427	3	US-09-173-581-10
9	291.4	13.9	1427	3	US-09-420-915-10
10	254	12.1	793	4	US-09-016-434-79
11	248	11.8	249	4	US-09-016-434-808
12	177.2	8.5	617	3	US-09-385-982-396
13	162.4	7.7	2100	1	US-08-264-002-6
14	155.4	7.4	1731	4	US-09-248-796A-6406
15	117	5.6	569	4	US-09-270-767-1008
16	117	5.6	569	4	US-09-270-767-16290
17	99.6	4.7	7218	1	US-08-232-463-14
18	97.6	4.7	326	4	US-09-513-999C-10592
19	91.4	4.4	641	4	US-09-270-767-1313
20	91.4	4.4	641	4	US-09-270-767-16595
21	62.6	3.0	891	4	US-09-248-796A-11247
22	60.6	2.9	51259	3	US-08-781-891-209
23	60.6	2.9	51259	4	US-09-618-166-209
24	59.8	2.9	1669	3	US-09-461-697-184
25	59.2	2.8	696	3	US-09-461-697-193
26	59.2	2.8	699	3	US-09-461-697-191
27	59.2	2.8	717	3	US-09-461-697-189

28	59.2	2.8	774	3	US-09-461-697-187	Sequence 187, App
29	59.2	2.8	819	3	US-09-461-697-185	Sequence 185, App
30	57.6	2.7	3211	2	US-08-574-959A-8	Sequence 8, Appli
31	57.6	2.7	3211	3	US-09-357-014-8	Sequence 8, Appli
32	57.6	2.7	3901	2	US-08-574-959A-6	Sequence 6, Appli
33	57.6	2.7	3901	3	US-09-357-014-6	Sequence 6, Appli
34	55.8	2.7	1434	4	US-09-248-796A-5873	Sequence 5873, Ap
35	55.8	2.7	2409	4	US-09-614-221A-204	Sequence 204, App
36	55.4	2.6	699	4	US-09-248-796A-9722	Sequence 9722, App
37	55	2.6	1737	4	US-09-248-796A-681	Sequence 681, App
38	54.8	2.6	478	4	US-09-621-976-10407	Sequence 10407, A
39	54.8	2.6	2223	1	US-08-257-073-4	Sequence 4, Appli
40	54.2	2.6	3337	1	US-08-072-610-1	Sequence 1, Appli
41	54.2	2.6	3337	2	US-08-719-822B-1	Sequence 1, Appli
42	54.2	2.6	3337	3	US-09-092-458-1	Sequence 1, Appli
43	54.2	2.6	3337	4	US-08-719-821A-103	Sequence 1, Appli
c	44	54	810	4	US-09-614-221A-103	Sequence 103, App
45	52.8	2.5	2518	3	US-09-433-699-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-759-359A-1
; Sequence 1, Application US/09759359A
; Patent No. 6492153
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDER, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/09/759,359A
; CURRENT FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3253
; TYPE: DNA
; ORGANISM: Human
US-09-759-359A-1

Query Match	100.0%	Score	2097	DB	4	Length	3253
Best Local Similarity	100.0%	Pred. No.	0				
Matches	2097	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	ATGAGCTCCCGAAAGTCTGCGCCATTTCAGGCCCGAAAGCGGAGCCCGAAAGAGAGAGAAA	60				
Db	91	ATGAGCTCCCGAAAGTCTGCGCCATTTCAGGCCCGAAAGCGGAGCCCGAAAGAGAGAGAAA	150				
Qy	61	CATCCGAAAAAGCCGAGGCTCAACAGAAAGCTCTTTAGTTTCTCTCTCTCTCCACCGCCA	120				
Db	151	CATCCGAAAAAGCCGAGGCTCAACAGAAAGCTCTTTAGTTTCTCTCTCTCTCCACCGCCA	210				
Qy	121	CCACACACACACCGCCACCTTTGCGAGACCCCAACCCCGGAGCCAGAGGAGAGATC	180				
Db	211	CCACACACACACCGCCACCTTTGCGAGACCCCAACCCCGGAGCCAGAGGAGAGATC	270				
Qy	181	CTGGGATCAGATGATGAGGAGCAAGGACCTTCGGGACTCTGCAAGGTGATATCAT	240				
Db	271	CTGGGATCAGATGATGAGGAGCAAGGACCTTCGGGACTCTGCAAGGTGATATCAT	330				
Qy	241	CCAGTGAATAATTTGGAGACCTTTCATGCGCGGTATCATGTTATTAGAAAGCTTGGATGG	300				
Db	331	CCAGTGAATAATTTGGAGACCTTTCATGCGCGGTATCATGTTATTAGAAAGCTTGGATGG	390				
Qy	301	GGGCACTTCTACTGTCTGGCTGTGCTGGGATATGACGGGAAAGATTTGTTGCAATG	360				
Db	391	GGGCACTTCTACTGTCTGGCTGTGCTGGGATATGACGGGAAAGATTTGTTGCAATG	450				
Qy	361	AAAGTTGTAAGTGGCCGACCATTTATACGAGACAGCCCTTGATCAATATAAATTTGCTC	420				

451 AAGTTGTAAGTGGCCAGCATTTATAGGAGACAGCCTTGGATGAAATAAAATTGCTC 510
Db
421 AAATGTGTCGAGAAAGTATCCAGTGCATCCCAAAACAAGACATGGTCCAGCTCAAT 480
Qy
511 AAATGTGTCGAGAAAGTATCCAGTGCATCCCAAAACAAGACATGGTCCAGCTCAAT 570
Db
481 GAGCACTTCAAGATTTCAAGCATGAATGGGATACATGTCTGATGGTCTTCAAGTACTT 540
Qy
571 GAGCACTTCAAGATTTCAAGCATGAATGGGATACATGTCTGATGGTCTTCAAGTACTT 630
Db
541 GGCACCACTCTCCTCAAGTGGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 600
Qy
631 GGCACCACTCTCCTCAAGTGGATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 690
Db
601 GTGAAGAGTATCATTTGACAGAGTCTTCAAGGTTTGAATTAATCAATCAATCAATCAAT 660
Qy
691 GTGAAGAGTATCATTTGACAGAGTCTTCAAGGTTTGAATTAATCAATCAATCAATCAAT 750
Db
661 ATCATTCACTGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 720
Qy
751 ATCATTCACTGACATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAAT 810
Db
721 AGAAGATGCGAGTCTGAGGCGCATGAGTGGGCAAGAGAGTGGTCTCTCTCTCTCTCTCT 780
Qy
811 AGAAGATGCGAGTCTGAGGCGCATGAGTGGGCAAGAGAGTGGTCTCTCTCTCTCTCTCT 870
Db
781 TCTGCACTGAGTACGCTCCAGCAGAGACCTATAGGAAATATCTAATAAACAAGAAAG 840
Qy
871 TCTGCACTGAGTACGCTCCAGCAGAGACCTATAGGAAATATCTAATAAACAAGAAAG 930
Db
841 AAAAACTGAAAGAAAGAGAGGCGAGGCTGAGTTATTGGAGAGGCGCTCGAGGAG 900
Qy
931 AAAAACTGAAAGAAAGAGAGGCGAGGCTGAGTTATTGGAGAGGCGCTCGAGGAG 990
Db
901 ATAGAAGATTTGAGCGAGAGAGTGAAGAGAAATTAATAGAGAGAAATCACTCAGCT 960
Qy
991 ATAGAAGATTTGAGCGAGAGAGTGAAGAGAAATTAATAGAGAGAAATCACTCAGCT 1050
Db
961 GCACCTTCAATGACAGAGTGGCAATCTGCGCAGAGTGAATCTAATAAACAAGAAAG 1020
Qy
1051 GCACCTTCAATGACAGAGTGGCAATCTGCGCAGAGTGAATCTAATAAACAAGAAAG 1110
Db
1021 TTAGAGAGGCGCTGAGGCGAGAGTGAAGAGAAATGGTGAAGTGAAGTGAAGTGAAGT 1080
Qy
1111 TTAGAGAGGCGCTGAGGCGAGAGTGAAGAGAAATGGTGAAGTGAAGTGAAGTGAAGT 1170
Db
1081 GAGAGAGAGATCTGAGAGAGAGAGATGAAGAGAAATGAAGATGAAGATGAAGATGAAG 1140
Qy
1171 GAGAGAGAGATCTGAGAGAGAGAGATGAAGAGAAATGAAGATGAAGATGAAGATGAAG 1230
Db
1141 CTTGCGAATGACACCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1200
Qy
1231 CTTGCGAATGACACCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1290
Db
1201 GGCCTTCTCTGAGGAGCACTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1260
Qy
1291 GGCCTTCTCTGAGGAGCACTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1350
Db
1261 AATCTGAGAGATTAATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1320
Qy
1351 AATCTGAGAGATTAATCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1410
Db
1321 TCCTATGAAATCAATGAGTGAATTCGCAATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1380
Qy
1411 TCCTATGAAATCAATGAGTGAATTCGCAATGAGAGTGAAGTGAAGTGAAGTGAAGTGAAG 1470
Db
1381 TTCCAGAGTTTCCACCTGTTTCTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1440
Qy
1471 TTCCAGAGTTTCCACCTGTTTCTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1530
Db
1441 GTGCTTCTGAGGATCACTTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1500
Qy
1531 GTGCTTCTGAGGATCACTTCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1590
Db

1501 AGCAGAACGGTTTTCAGCTCCAGTACTGGGATTTGCAAAAGCAAAACCCGGCAGCT 1560
Qy
1591 AGCAGAACGGTTTTCAGCTCCAGTACTGGGATTTGCAAAAGCAAAACCCGGCAGCT 1650
Db
1561 GACTTGTGTGTAATCCCTGGATCCGCGGAATGCAAGATTAATAATAGAGTAAATAATGCT 1620
Qy
1651 GACTTGTGTGTAATCCCTGGATCCGCGGAATGCAAGATTAATAATAGAGTAAATAATGCT 1710
Db
1621 GACCTGGGAATGCTTGTGGTGTCATAAACACATTCACGGAAGACATCCAGACCGCTCAG 1680
Qy
1711 GACCTGGGAATGCTTGTGGTGTCATAAACACATTCACGGAAGACATCCAGACCGCTCAG 1770
Db
1681 TACCGCTCCATAGAGGTTTAAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGGAGC 1740
Qy
1771 TACCGCTCCATAGAGGTTTAAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGGAGC 1830
Db
1741 ACGCGGTGTATGGATTTGAGCTGGCAACCGGAGATTAATTTGTTGAAACCAATCTCTGGG 1800
Qy
1831 ACGCGGTGTATGGATTTGAGCTGGCAACCGGAGATTAATTTGTTGAAACCAATCTCTGGG 1890
Db
1801 GAAGACTATTCCAGAGACGAGACCAATAGCCCAATCATAGAGCTGCTAGGCAATTT 1860
Qy
1891 GAAGACTATTCCAGAGACGAGACCAATAGCCCAATCATAGAGCTGCTAGGCAATTT 1950
Db
1861 CCAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTTCAATCGCAGAGGAA 1920
Qy
1951 CCAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTTCAATCGCAGAGGAA 2010
Db
1921 CTGGACACATCAACAGCTGAGCCCTGGAGCCTTTGATGATCTTGTGGAAGATTTAT 1980
Qy
2011 CTGGACACATCAACAGCTGAGCCCTGGAGCCTTTGATGATCTTGTGGAAGATTTAT 2070
Db
1981 GCCTGGCCCATGAGATGCTGCACAGTTTACAGATTTCTGATCCCGATTTAGAAATG 2040
Qy
2071 GCCTGGCCCATGAGATGCTGCACAGTTTACAGATTTCTGATCCCGATTTAGAAATG 2130
Db
2041 GTTCAGAAAAACGAGCCTCAGCTGGGGAATGCTTGGGATCTCTTGGTGAATTTCT 2097
Qy
2131 GTTCAGAAAAACGAGCCTCAGCTGGGGAATGCTTGGGATCTCTTGGTGAATTTCT 2187
Db

RESULT 2

US-10-207-973-1
; Sequence 1, Application US/10207973
; Patent No. 6753175
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/10/207,973
; CURRENT FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3253
; TYPE: DNA
; ORGANISM: Human
US-10-207-973-1

Query Match 100.0%; Score 2097; DB 4; Length 3253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAGCTCCCGAAAGTCTGGCCATTTCAGGCCGCAAGAGCGAGGCCGCAAAAGAGAGAAA 60
Db 91 ATGAGCTCCCGAAAGTCTGGCCATTTCAGGCCGCAAGAGCGAGGCCGCAAAAGAGAGAAA 150
Qy 61 CATCCGAAAAAGCCGAGGCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCTCTCTCTCTCT 120
Db 151 CATCCGAAAAAGCCGAGGCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCTCTCTCTCTCT 210

;; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
;; NUMBER OF SEQUENCES: 12
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Drive
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/016,000
;; FILING DATE: HERewith
;; CLASSIFICATION:
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0465 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-855-0555
;; TELEFAX: 650-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 8:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 2791 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: HEARNOT01
;; CLONE: 307624
;; US-09-016-000-8

Query Match 96.5%; Score 2023.8; DB 2; Length 2791;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2028; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 63 TCCGAAAAGCGGAGCCTCAACAGAAAGCTCCTTTAGTTTCCTCTCTCCCTCCACGCCACC 122
DB 202 TTCAGAAAGCGCGAGCCTCAACAGAAAGCTCCTTTAGTTTCCTCTCTCCCTCCACGCCACC 261
QY 123 ACCACACACCGCCACCTTTTCCGACAGACCCACACCCCGGAGCCAGAGGAGATCCT 182
DB 262 ACCACACACCGCCCGCTTTTCCGACAGACCCACACCCCGGAGCCAGAGGAGATCCT 321
QY 183 GGGATCAGATGATGAGGACGAGGAGCCCTCGGACTACTCGAAGGTGGATATCATCC 242
DB 322 GGGATCAGATGATGAGGACGAGGAGCCCTCGGACTACTCGAAGGTGGATATCATCC 381
QY 243 AGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTTGGATGGG 302
DB 382 AGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTTGGATGGG 441
QY 303 GCATTTCTCTACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 362
DB 442 GCATTTCTCTACTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 501
QY 363 AGTGTAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAAATAAAATTTGCTCAA 422
DB 502 AGTGTAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAAATAAAATTTGCTCAA 561
QY 423 ATGTGTTTCAGAAAGTGATCCAGTGACCCCAACCAAGACATGTTGGTCCAGCTCATTTGA 482
DB 562 ATGTGTTTCAGAAAGTGATCCAGTGACCCCAACCAAGACATGTTGGTCCAGCTCATTTGA 621
QY 483 CGACTTCAAGATTTACGGATGATGGGATACATGCTGCTGCTGCTTCCGAGTACTTGG 542

DB 622 CGACTTCAAGATTTACGGCATGAATGGGATACATGCTGATGGTCTTCGAGTACTTGG 681
QY 543 CCACCATCTCCTCAAGTGGATCATCAATCAACTATCAAGGCCTCCAGTACCTTGGT 602
DB 682 CCACCATCTCCTCAAGTGGATCATCAATCAACTATCAAGGCCTCCAGTACCTTGGT 741
QY 603 GAAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTATAGTCAAGAT 662
DB 742 GAAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTATAGTCAAGAT 801
QY 663 CATTCACTACTGACATAAAGCCGGAATAATCTTCTGATGTGTGTGATGATGATGATGATGATGAT 722
DB 802 CATTCACTACTGACATAAAGCCGGAATAATCTTCTGATGTGTGTGATGATGATGATGATGATGAT 861
QY 723 AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCCTCTCTCAGGTC 782
DB 862 AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCCTCTCTCAGGTC 921
QY 783 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTTATAGGAAAATATCTTAAACCAAAAGAA 842
DB 922 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTTATAGGAAAATATCTTAAACCAAAAGAA 981
QY 843 AAAACTGAAAAGAAACAGAGAGCGCAGGCTGAGTATTGGAGAAAGCGCCTGCAGGAGAT 902
DB 982 AAAACTGAAAAGAAACAGAGAGCGCAGGCTGAGTATTGGAGAAAGCGCCTGCAGGAGAT 1041
QY 903 AGAAGAATTTGAGCGGAGAGCTGAAAGGAAAATTAATAGAGAAAACATCAGCTCAGCTGC 962
DB 1042 AGAAGAATTTGAGCGGAGAGCTGAAAGGAAAATTAATAGAGAAAACATCAGCTCAGCTGC 1101
QY 963 ACCTTCCAAATGACCCAGGATGGCGAATCTGCTCCAGAGGTGAACTTAAACCAACAGGATT 1022
DB 1102 ACCTTCCAAATGACCCAGGATGGCGAATCTGCTCCAGAGGTGAACTTAAACCAACAGGATT 1161
QY 1023 AGAGGAGCGGCTGAGCGAGAGCTGCAAGAGCAATGCTGAGCTGAGGACCCAGGAAGA 1082
DB 1162 AGAGGAGCGGCTGAGCGAGAGCTGCAAGAGCAATGCTGAGCTGAGGACCCAGGAAGA 1221
QY 1083 GAAAGAGATGCTGAGAAAGAAAACATTGAAAAAGATGAAGATGATGATGATGATGATGATGATGAT 1142
DB 1222 GAAAGAGATGCTGAGAAAGAAAACATTGAAAAAGATGAAGATGATGATGATGATGATGATGATGAT 1281
QY 1143 TCGGAAATAGACCCCTAGCTGATAGATCACTTAAACCAATGCTGATGATGATGATGATGATGATGATGAT 1202
DB 1282 TCGGAAATAGACCCCTAGCTGATAGATCACTTAAACCAATGCTGATGATGATGATGATGATGATGATGAT 1341
QY 1203 CCCATTCTCACTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAGACTGCCCAAA 1262
DB 1342 CCCATTCTCACTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAGACTGCCCAAA 1401
QY 1263 TCCTGAGGAATATATCTTGTATGAGCCAAATGCGAGAAAGTGAATTAACATATAGCAGCTC 1322
DB 1402 TCCTGAGGAATATATCTTGTATGAGCCAAATGCGAGAAAGTGAATTAACATATAGCAGCTC 1461
QY 1323 CTATGAACAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1382
DB 1462 CTATGAACAAATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1521
QY 1383 CCCAGATTTTCCACCTCGTTGTTCTCTGAGATGCTTGAACCTGTGGCTGCGGCTCTGT 1442
DB 1522 CCCAGATTTTCCACCTCGTTGTTCTCTGAGATGCTTGAACCTGTGGCTGCGGCTCTGT 1581
QY 1443 GCTTTTTCAGGGAATCACCCTTACTGTAGCAAGAGAGAGAGTCCATCCCATGACAGAAG 1502
DB 1582 GCTTTTTCAGGGAATCACCCTTACTGTAGCAAGAGAGAGAGTCCATCCCATGACAGAAG 1641
QY 1503 CAGAACGTTTTCAGGCTCCAGTCTGAGGATTTGCCAAAGCAAAACCCCGGAGCTGA 1562
DB 1642 CAGAACGTTTTCAGGCTCCAGTCTGAGGATTTGCCAAAGCAAAACCCCGGAGCTGA 1701
QY 1563 CTGTGTTGTTGAATCCCTCGGATGCCGGAATGCGATAAATTTAGAGTAAATTTGCTGA 1622

Db 63657 TTGTTCTCTGGATCCTTAGAACCTGTGGCTGCGGCTCTGTGCTTTCTGAGGGATCACCA 63716
QY 1462 CTTACTGAGCAGAGGAGCAGTCCATCCATGACAGAAAGCAGAACGGTTTTCAGCCTCC 1521
Db 63717 CTTACTGAGCAGAGGAGCAGTCCATCCATGACAGAAAGCAGAACGGTTTTCAGCCTCC 63776
QY 1522 AGTACTGGGGATTTCGCAAAAGCAAA 1547
Db 63777 AGTACTGGGGATTTCGCAAAAGGTAA 63802

RESULT 8

US-09-173-581-10
; Sequence 10, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 119819
US-09-173-581-10

Query Match 13.9%; Score 291.4; DB 3; Length 1427;

Best Local Similarity 70.2%; Pred. No. 1.1e-68;
Matches 391; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 1534 TTGCAAAAGCAAAACCCGGGCAGCTGACTTTGGTGAATCCCTCGATCCGGGAAT 1593
Db 800 TCGCCTAGCACACCATTCGGTGCCTCGAACCTCTGGTGAACCCCTCGAGCCCCAAAT 859
QY 1594 GCAGATAAAATTAGAGTAAATTTGCTGACCTGGGAATGCTTTGGTGCATAAACAC 1653
Db 860 GCAGATAAGATCAAGATCAAGATCGAGACCTGGGCACGCTGCTGGGTGCACAAGCAC 919
QY 1654 TTCAGGAAGACATCCAGACGGTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGG 1713
Db 920 TTCAGGAAGACATCCAGACTCGGAGTACCGGGCGCTCGAGTGTGATCGGCCGCA 979
QY 1714 TACAGCACCCCTCGGACATCTGGAGCAGCGGTGTATGGCAATTTGAGCTGCAACGGGA 1773
Db 980 TACGGCCCCCGGACACATCTGGAGCAGCAGCTGATGGCTTCGAGCTGGCCACTGCT 1039
QY 1774 GATTATTTGTTGAACCAATTTCTGGGGAAGACTTTCAGAGACGAAGACACATAGCC 1833
Db 1040 GACTACCTGTTGAGCGCGATTTCTGGAGAAGACTACAGTCGTGATGAGGACCAATCGCT 1099
QY 1834 CACATCATAGAGCTGTAGCGAGTATCCAGGCACTTTTGTCTATCTCGAAATATTTCT 1893
Db 1100 CACATAGTGGAGCTTCTGGGGACATTCGCCCGCTTCGCCCTCTCAGGCCGCTATTCC 1159
QY 1894 CGGGAAATTTCTCAATCGCAGAGAGAACTCGGACACATCACCAAGCTGAAGCCCTGGAGC 1953
Db 1160 CGGGAGTTCTTCAACCGGAGAGAGAGCTCGGCAATCCCAATCTCAAGCACTGGGCG 1219
QY 1954 CTCCTTGATGTTCTGTGAAAAGTATGGCTGGCCCATGAAGATGCTGCACAGTTTACA 2013

Db 1220 CTGTACGAGGTACTCATGTGAAAAGTACGAGTGGGCCCTTAGAGCGGCACACAGTTTCAGC 1279
QY 2014 GATTTCCTGATCCCGATGTTAGAAATGTTCCAGAAAAACGAGCCTCAGCTGGCGAATGC 2073
Db 1280 GCCTTTCTGCTGCCATGAATGAGTACATCCCCGAAAAGCGGCCAGTCCCGTCACTGC 1339
QY 2074 CTTGGGCATCCTTGTT 2090
Db 1340 CTCCAGCACCCCTGGCT 1356

RESULT 9

US-09-420-915-10
; Sequence 10, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/420,915
; EARLIER FILING DATE: 1999-10-20
; EARLIER APPLICATION NUMBER: US 09/173,581
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: -
; OTHER INFORMATION: 119819
US-09-420-915-10

Query Match 13.9%; Score 291.4; DB 3; Length 1427;

Best Local Similarity 70.2%; Pred. No. 1.1e-68;
Matches 391; Conservative 0; Mismatches 166; Indels 0; Gaps 0;

QY 1534 TTGCAAAAGCAAAACCCGGGCAGCTGACTTTGGTGAATCCCTCGATCCGGGAAT 1593
Db 800 TCGCCTAGCACACCATTCGGTGCCTCGAACCTCTGGTGAACCCCTCGAGCCCCAAAT 859
QY 1594 GCAGATAAAATTAGAGTAAATTTGCTGACCTGGGAATGCTTTGGTGCATAAACAC 1653
Db 860 GCAGATAAGATCAAGATCAAGATCGAGACCTGGGCACGCTGCTGGGTGCACAAGCAC 919
QY 1654 TTCAGGAAGACATCCAGACGGTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGG 1713
Db 920 TTCAGGAAGACATCCAGACTCGGAGTACCGGGCGCTCGAGTGTGATCGGCCGCA 979
QY 1714 TACAGCACCCCTCGGACATCTGGAGCAGCGGTGTATGGCAATTTGAGCTGCAACGGGA 1773
Db 980 TACGGCCCCCGGACACATCTGGAGCAGCAGCTGATGGCTTCGAGCTGGCCACTGCT 1039
QY 1774 GATTATTTGTTGAACCAATTTCTGGGGAAGACTTTCAGAGACGAAGACACATAGCC 1833
Db 1040 GACTACCTGTTGAGCGCGATTTCTGGAGAAGACTACAGTCGTGATGAGGACCAATCGCT 1099
QY 1834 CACATCATAGAGCTGTAGCGAGTATTCAGGCACTTTTGTCTATCTCGAAATATTTCT 1893
Db 1100 CACATAGTGGAGCTTCTGGGGACATTCGCCCGCTTCGCCCTCTCAGGCCGCTATTCC 1159
QY 1894 CGGGAAATTTCTCAATCGCAGAGAGAACTCGGACACATCACCAAGCTGAAGCCCTGGAGC 1953
Db 1160 CGGGAGTTCTTCAACCGGAGAGAGAGCTCGGCAATCCCAATCTCAAGCACTGGGCG 1219
QY 1954 CTCCTTGATGTTCTGTGAAAAGTATGGCTGGCCCATGAAGATGCTGCACAGTTTACA 1219

Qy	1954	CTCTTTGATCTACTGTGGAAAAGTATGGCTGGCCCCATGAAGTCTGTCACAGTTTACA	2013
Db	1220	CTGTACGAGTACTCATCTGGAAGAACTACGATGGCCCCCTAGACAGGCCACACAGTTTCAGC	1279
Qy	2014	GATTTCTGTATCCGATGTTTAGAAATGGTTCCAGAAAAACGAGCTCAGCTGGCGGAATGC	2073
Db	1280	GCCTTTTCTGTCGCCATGAATGAGTATACCCCGAAAAGGGGCCAGTGCCTCGTACTGC	1339
Qy	2074	CTTCGGCATCCTTGGTT	2090
Db	1340	CTCAGACACCCCTGGCT	1356

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

	Query Match Best Local Similarity Matches	12.1%; 67.4%; 360;	Score 254; Pred. No. 1.1e-58; 0;	DB 4; 1.1e-58; Mismatches 172;	Length 793; Indels 2; Gaps 2
Qy	179	TCCTGGGATCAGATGATGAGGACGAGGACCTTCGCGGACTACTGCAAGGTGGATATC	238		
Db	156	TTCTGGGCTCCGACGACGAGGAAACAGGGAAGACCCCAAGAGACTACTGCAAGGCGCGCTACC	215		
Qy	239	ATCCAGTGAATAATTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGAT	298		
Db	216	ACCCGTGAAGATCGCGGACCTGTTCAATGGCGGTACACGTTGCTGNCNCNCNCNNGT	275		

299	Qy	GGGGGCACTCTCTCTACGTCTGGCTGTGCTGGATATGACAGGGGAAAAGATTGTGTGGAA	358
276	Db	GGGGCCACTCTTCCACCGTCTGGCTCTGCTGGGANTTCNNCGCAAGCGCTTGTGGCCC	335
359	Qy	TGAAAGTTGTAAAAAGTCCACGATTTATAC - GGAGACAGCGCTTGGATGAATAAAATTG	417
336	Db	TCAAGTGGTGAANAGTGGGGGCAITTCACGGGAGACAGCTGTGGATGAATCAANCTC	395
418	Qy	CT - CAAATGTGTCGAGAAAGTGATCCAGGTGACCCAAACAAAGACATGTGTGGTCCAGCT	476
396	Db	CTGAAATGTGTCGGGACACGACCCAGTGACCCCAAPAGAAACCATTTGTCCANCT	455
477	Qy	CATTGAGCACTTCAAGATTTTCAGGCATGAATGGGATACATGCTCTGCATGTGCTTCCGAAGT	536
456	Db	CATTGTATGACTTCAGGATCTCANGAGTCAATGGGANTCCAATGTGTGCATGTGTGGANGT	515
537	Qy	ACTTTGGCCACCATCTCCTCAAGTGGATCATCAATCCAATCATCAAGGCGCTCCCAAGTACG	596
516	Db	GCTGGGGCACCANCTCCTCAAAATGGATGTCNTCAANTCCAATCAACGCGCTGCCCGTGCC	575
597	Qy	TTGTGTGAAGAGTATCATTCGACAGGTCCTTCAAGGGTTAGATTACTTTACACAGTAAAGTG	656
576	Db	CTGCCITTAAANAACATCNTNWANGANGTGCTGCACNGCTGNANTACCCCAACCNANTTG	635
657	Qy	CAAGATCATTTACACTGACATAAAGCCGGAAAAATATCTTTGATGTGTGTGGATGA	710
636	Db	CAANATCNTCCANCNGGAATCAANCCCAAAACATCTTGTGTGTGTGGGGGA	689

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

LIBRARY: BRAINOT03
CLONE: 529450
US-09-016-434-808

Query Match 11.8%; Score 248; DB 4; Length 249;
Best Local Similarity 99.6%; Pred. No. 2.3e-57;
Matches 248; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1511 TTTCAGCTCCAGTACTGGGATTTGCCAAAGCAAAAACCGGGCAGCTGACTTTGG 1570
DB 1 TTTCAGCTCCAGTACTGGGATTTGCCAAAGCAAAAACCGGGCAGCTGACTTTGG 60

QY 1571 TGAATCCCTCGATCCGGGATGAGATATAAATTAGAGTAAATTTCTGACCTGGGAA 1630
DB 61 TGAATCCCTCGATCCGGGATGAGATATAAATTAGAGTAAATTTCTGACCTGGGAA 120

QY 1631 ATGCTTTGTTGGTGCATTAACACTTTCAGGAGACATCCAGACGGCTCAGTACCGCTCCA 1690
DB 121 ATGCTTTGTTGGTGCATTAACACTTTCAGGAGACATCCAGACGGCTCAGTACCGCTCCA 180

QY 1691 TAGAGTTTAAATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAGCAGGGGTGTA 1750
DB 181 TAGAGTTTAAATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAGCAGGGGTGTA 240

QY 1751 TGGCATTTG 1759
DB 241 TGGCATTTG 249

RESULT 12
US-09-385-982-396
Sequence 396, Application US/09385982
Patent No. 6262334
GENERAL INFORMATION:
APPLICANT: ENDEGE, WILSON O., ET AL.
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCDNA-260XX
CURRENT APPLICATION NUMBER: US/09/385,982
CURRENT FILING DATE: 1999-08-30
EARLIER APPLICATION NUMBER: 09/328,111
EARLIER FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: 60/117,393
EARLIER FILING DATE: 1999-01-27
EARLIER APPLICATION NUMBER: 60/098,639
EARLIER FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 544
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 396
LENGTH: 617
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(617)
OTHER INFORMATION: n = A,T,C or G

US-09-385-982-396

Query Match 8.5%; Score 177.2; DB 3; Length 617;
Best Local Similarity 71.1%; Pred. No. 5.4e-36;
Matches 275; Conservative 0; Mismatches 109; Indels 3; Gaps 3;

QY 160 CCGGAGCCAGAGGAGGATCTCTGGGATCAGATGATGAGGAGCAAGAGGACCTCGGGAC 219
DB 177 CCAGAGCAGGAGAGGAGATCTCTGGGATCTGATGATGATGAGGAGCAAGATCTAATGAT 236

QY 220 TACTGCAAGGTGGATATCATTCAGTAAATTTGAGACCTCTTCAATGCGCGGTATCAT 279
DB 237 TATTGTAAGGAGGTATCATCTTGTGAAAAATTGGAGATCTATTTCATATGGAGATACCAT 296

QY 280 GTTATTAGAACTTTGGATGGGCACTTCTCTACTGCTGGCTGTCTGGGATATGTCAG 339
DB 297 GTGATCCGAAAGTTAGGCTGGGACACTTTTCAACAGTATGGTTATCATGGGATATTCAG 356

QY 340 GCGAAAGATTTCGAATGAAGTTGTAAAAAGTCCCGCAGCATTTATAC-GGAGACAGC 398
DB 357 GCGAAGAAATTTCTGGCAATGAAGTAGTTAAAGTGTCTGAACATTACACTTGAACCCAG 416

QY 399 CTTGGATGAATAAATTTGCTCAAAATGTTTCAGAAAGTGTCCCGAGTACCCAGTACCCAAACAA 458
DB 417 CTTAGATGAATCCGGTTGCTTGAAGTCAAGTTTCCAATTCAGACCTTATGGATCCAAATNG 476

QY 459 AGACATGGTGGTCCAGCTCATTCAGCACTTCAAGATTTCAGGCATGAATGGGATACATGT 518
DB 477 AAAAATGGTTGT-CAACTACTAGTACITTA- AATTAGGAGTTAATGGAACACATAT 534

QY 519 CTGCATGCTCTTCGAAAGTACTTGGCCA 545
DB 535 TTGCATGGATTTGAAGTTTGGGCA 561

RESULT 13
US-08-264-002-6
Sequence 6, Application US/08264002
Patent No. 5559019
GENERAL INFORMATION:
APPLICANT: GUI, JIAN-FANG
APPLICANT: FU, XIANG-DONG
TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
STREET: 1880 Century Park East, Fifth Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90067

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/264,002
FILING DATE: 22-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: TUMARKIN PH.D., LISA A.
REGISTRATION NUMBER: P-38,347
REFERENCE/DOCKET NUMBER: PD3590
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/455-5100
TELEFAX: 619/455-5110
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2100 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: DSK1
FEATURE:
NAME/KEY: CDS
LOCATION: 224...1855

US-08-264-002-6

Query Match 7.7%; Score 162.4; DB 1; Length 2100;
Best Local Similarity 58.1%; Pred. No. 1.1e-33;
Matches 286; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

QY 1600 AAAATTAGATGAATAAATTTCTGACCTGGGAATGCTTGTGGTGCATATAACACTTCACG 1659
DB 1280 AAAATCACTGTAAAAATTTGCCGATCTCGGTAAACGGTCTCGGACACAGGAGATTTCACC 1339

QY 1660 GAAGACATCCAGACGGCTCAGTACCGCTCCATAGAGTTTAAATAGGAGGGGTACAGC 1719

Db 1340 AACGATGTCNAACCGCTAGTATAGTCTCCAGAAGTAATTCCTAGATGTCCTCGGGA 1399
Qy 1720 ACCCTGCGGACATCTGGAGCAGCGGTGATGGCAATTTGAGCTGGCAACGGGAGATTAT 1779
Db 1400 GCTTCGCTGATGCTGGAGTTTGGCTGTATCATTTTGAATTTGCTAACCGCGGATTAC 1459
Qy 1780 TTGTTTGAACCAATCTTGGGGAAGACTATTCAGAGACGAACACATAGCCACATC 1839
Db 1460 CTTTTCGATCCCGAATGGGAATTTCTATCTTAAGGAGGATGACCAACATGCCCCAAAT 1519
Qy 1840 ATAGAGCTGTAGCGAGTATTCAGGCACTTTGGCTCTATCTGGAATAATTCCTCGGGA 1899
Db 1520 ATTGAGTATTGTTTAATATCTTAAGCAATGGCACTTTCAGAAAGCACTCCCCGAT 1579
Qy 1900 TTCTTCAATCGCAGAGAGAACTTCGCACACATCAACAGCTGAAGCCCTGGAGCTCTTT 1959
Db 1580 TTATTTAAACCGTCGCGTGAACCTCCGAATATTCATAAATTTGAAGTTTGGCCCTTAAAA 1639
Qy 1960 GATGTACTTGTGAAAGTATGCTGGCCCCATGAAGATGCTGCAGATTACAGATTTC 2019
Db 1640 GATGTTTGGAGCAAAATACCAATTTTTCAGCCGAGTTGGCTCAACAAATATCAGACTTT 1699
Qy 2020 CTGATCCGATGTAGAAATGGTTCCAGAAACGAGCCTCAGCTGCGAATGCCCTTCG 2079
Db 1700 TTATCTCTATGTTATGTTTGTATCTGCGAAGCAACCAATGCTGGTTACATGACCAAT 1759
Qy 2080 CATCCTTGGTTG 2091
Db 1760 TCTCCATGGTTG 1771

RESULT 14

US-09-248-796A-6406
; Sequence 6406, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 6406
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6406

Query Match 7.4%; Score 155.4; DB 4; Length 1731;
Best Local Similarity 56.9%; Pred. No. 7.8e-32;
Matches 285; Conservative 0; Mismatches 216; Indels 0; Gaps 0;
Qy 1592 ATGCAGATAAAATTAGAGTAAATAATTCGACCTGGGAATGCTTGGGTGCATAAATAC 1651
Db 980 ATAATGAATTAATATCTGTTAAATTTGCTGATTTGGGTAAATGATGTTGACAAATCATC 1039
Qy 1652 ACTTCAAGGAACATCCAGACGCGTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGG 1711
Db 1040 ATTTTACTGATGAATCCAAACAAAGCAATATCGAGCCCCGAAAAATTTAAATGGTTATT 1099
Qy 1712 GGTAGACGCCCTCGGACATCTGGAGCAGCGGTGATGGCAATTTGAGCTGGCAACGG 1771
Db 1100 ATTGGGGTGCCTGAGTGAATTTATGCTCAATTTGCCCTGTTAATATTTGAATTTAACTG 1159
Qy 1772 GAGATTATTGTTTGAACACATTTCTGGGGAGACATTTCCAGAGACGAGACCACATAG 1831
Db 1160 GTGATTATCTATTGATCTCTAGAGATGGTAAATCATATATAAAAAGATGATGATCATATTG 1219

Qy 1832 CCCACATCATAGAGCTGTAGGCAAGTATTTCCAAGGCACTTTTGCTCTATCTGGAATAATTT 1891
Db 1220 CCCAAATTATTGAATTTGATTTGGACCAATTTCTCAATCAAAATGTTAAAAAGAAAGTTATTATG 1279
Qy 1892 CTCGGGAATTTCTCAATCCAGAGAGAACTCGGCACACATCACCAGCTGAAGCCCTGGA 1951
Db 1280 CTCGAGAAATTTTCAATTCAGGTATGAATTAAGAAGAATCATGAATTTAAACCTTTGGG 1339
Qy 1952 GCCTCTTTGATGTACTTTGGAAAAGTATGGCTGGCCCCCATGAAGATGCTGCACAGTTTA 2011
Db 1340 GGTTCACAGATGTAATTAATTTGAAAATATAAATTCCTTTGATGATGCTATTGAAATTA 1399
Qy 2012 CAGATTTCTGATCCCGATGTTAGAAATGGTTCCAGAAAAACGAGCTCAGCTGGCGAAT 2071
Db 1400 GTGAATTTTATTTGCCAATGTTGAAATTTAAACACAGAGAAGAGCTGATGCTGGAGAA 1459
Qy 2072 GCCTTCGCGATCCTTGGTTGA 2092
Db 1460 TGTGTAATCATCTTGGTTAA 1480

RESULT 15

US-09-270-767-1008
; Sequence 1008, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1008
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1008

Query Match 5.6%; Score 117; DB 4; Length 569;
Best Local Similarity 67.3%; Pred. No. 9.9e-22;
Matches 165; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
Qy 1600 AAAAATTAGAGTAAAAATTTGCTGACCTGGGAATGCTTGTGGTGCATAAAACACTTCACG 1659
Db 324 AACGTTGGGTAAAGATCGCCGACTTGGGGAACGCTGCTACGACTACCATCACTTTACT 383
Qy 1660 GAAGACATCCAGACGGTCACTACCGCTCCATAGAGTTTAAATAGAGGGGTACAGC 1719
Db 384 GAGGACATTCAGACTCGCCAGTATCGATCAATCGAGGTTCTTTTGGGAGCGCGTACAAT 443
Qy 1720 ACCCTGCGGACATCTGGAGACGCGGTGATGGCAATTTGAGCTGCAACGGGAGATTAT 1779
Db 444 TATACCGCGACATCTGGAGCAGACGCTGTTGGCTTCGAGCTGCGCCACCGCGACTAC 503
Qy 1780 TTGTTTGAACCACTTCTGGGAAGACTATTTCAGAGAGAAAGACCAACATAGCCCAATC 1839
Db 504 CTGTTGACCCCTCACGCCGAGAGTCTTACAGTCGGGACGAGGACCACTTGGCGCATATT 563
Qy 1840 ATAGA 1844
Db 564 GTGGA 568

Search completed: January 23, 2005, 11:11:34
Job time : 166.232 secs

11113 ruyg blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2005, 05:50:23 ; Search time 8680.01 Seconds
(without alignments)
11424.696 Million cell updates/sec

Title: US-10-799-676-1_COPY_91_2187
Perfect score: 2097
Sequence: 1 atgagctccggaaagtgc.....ggcatccttggtgaattct 2097

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2097	100.0	3253	6	AR265348 Sequence
2	2097	100.0	3253	6	AX666205 Sequence
3	2028.6	96.7	3157	9	BC035214 Homo sapi
4	2023.8	96.5	2577	9	BC068547 Homo sapi
5	2023.8	96.5	2791	6	AR077238 Sequence
6	2023.8	96.5	2791	6	BD107902 Protein k
7	2006.6	95.7	3708	6	CQ715987 Sequence
8	1991.8	95.0	3745	6	AX429244 Sequence
9	1991.8	95.0	3745	6	AX666550 Sequence
10	1991.8	95.0	3745	9	HSU88666 Homo sapien
11	1814.6	86.5	3028	9	AY354201 Homo sapi
12	1662.6	79.3	2300	10	AB006036 Mus muscu
13	1652.2	79.0	3578	10	BC020178 Mus muscu
14	1652.2	78.8	4287	10	BC062941 Mus muscu
15	1284	61.2	2565	9	AK093420 Homo sapi
16	969.2	46.2	2326	9	AK090783 Homo sapi
17	833.6	39.8	183097	9	AC087358 Homo sapi
18	830	39.6	149318	9	AC091554 Homo sapi
19	828.4	39.5	183097	9	AC087358 Homo sapi

C	20	828.4	39.5	183360	2	AC022525	AC022525 Homo sapi
C	21	827.6	39.5	149318	9	AC091554	AC091554 Homo sapi
	22	821.6	39.2	183360	2	AC022525	AC022525 Homo sapi
	23	723.4	34.5	917	10	MMU92456	U92456 Mus musculu
	24	677.4	32.3	984	6	CQ739202	CQ739202 Sequence
	25	631.4	30.1	4299	6	I26663	I26663 Sequence 1
	26	631.4	30.1	4326	6	AX410648	AX410648 Sequence
	27	631.4	30.1	4326	6	AX666548	AX666548 Sequence
	28	631.4	30.1	4326	9	HSU09564	U09564 Human serin
	29	631.4	30.1	4760	9	HSU09564	AJ318054 Homo sapi
	30	629.4	30.0	923	6	AR270140	AR270140 Sequence
	31	628.2	30.0	4252	6	CQ714961	CQ714961 Sequence
	32	628.2	30.0	4639	6	CQ414916	CQ414916 Sequence
	33	626.6	29.9	4296	9	HSU808814	EX648663 Homo sapi
	34	626.6	29.9	4297	9	HSU808848	EX648697 Homo sapi
	35	626.6	29.9	4388	9	BC038292	BC038292 Homo sapi
	36	621.6	29.6	2460	10	BC005707	BC005707 Mus muscu
	37	621.6	29.6	2485	10	BC050761	BC050761 Mus muscu
	38	620	29.6	2416	10	AB012290	AB012290 Mus muscu
	39	615.2	29.3	2469	10	MMAJ4115	AJ224115 Mus muscu
	40	607.6	29.0	4071	10	AF446079	AF446079 Cricetulu
	41	565.4	27.0	984	5	CR385926	CR385926 Gallus ga
	42	563.8	26.9	2628	5	BC056825	BC056825 Danio rer
	43	486.8	23.2	90541	6	AR265349	AR265349 Sequence
	44	486.8	23.2	90541	6	AX666207	AX666207 Sequence
C	45	486.8	23.2	113033	9	AC005070	AC005070 Homo sapi

ALIGNMENTS

RESULT 1
AR265348
LOCUS AR265348
DEFINITION Sequence 1 from patent US 6492153.
ACCESSION AR265348
VERSION AR265348.1 GI:296939850
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3253)
AUTHORS Abu-Threideh,J., Gong,F., Ketchum,K.A., Di Francesco,V. and Beasley,E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6492153-A 1 10-DEC-2002;
FEATURES
Location/Qualifiers
source 1..3253
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN	Query Match	100.0%;	Score 2097;	DB 6;	Length 3253;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2097;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGAGTCTCCGAAAAGTCTGTCGCCATTTCAGGCCCGAAAGCGGAGCCGAAAGAGAGAAA	60		
DB	91	ATGAGTCTCCGAAAAGTCTGTCGCCATTTCAGGCCCGAAAGCGGAGCCGAAAGAGAGAAA	150		
QY	61	CATCCGAAAAACCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCTCTCCACCGCCA	120		
DB	151	CATCCGAAAAACCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCTCTCCACCGCCA	210		
QY	121	CCACCACCAACCGCCACCTTTTGGCAGACCCACACACCCCGGAGCCAGGAGGAGATC	180		
DB	211	CCACCACCAACCGCCACCTTTTGGCAGACCCACACACCCCGGAGCCAGGAGGAGATC	270		
QY	181	CTGGGATCAGATGATGAGGAGCAAGAGGACCTTCGGGACTACTCTGCAAGGTGATATCAT	240		
DB	271	CTGGGATCAGATGATGAGGAGCAAGAGGACCTTCGGGACTACTCTGCAAGGTGATATCAT	330		

QY 241 CCAGTGAAATTTGGAGACCTCTTCAATGCGCGGTATCATGTTATTAGAAAAGCTGGATGG 300
Db 331 CCAGTGAAATTTGGAGACCTCTTCAATGCGCGGTATCATGTTATTAGAAAAGCTGGATGG 390
QY 301 GGGCACTTCTTCTACTGTCTGTGCTGTGCGGATATGCGAGGGGAAAGATTGTTGCAATG 360
Db 391 GGGCACTTCTTCTACTGTCTGTGCTGTGCGGATATGCGAGGGGAAAGATTGTTGCAATG 450
QY 361 AAAGTTGTAAGAAGTCCAGCATTTATACGGAGACAGCTTGGATGAATAAATAATTTGCTC 420
Db 451 AAAGTTGTAAGAAGTCCAGCATTTATACGGAGACAGCTTGGATGAATAAATAATTTGCTC 510
QY 421 AAATGTGTTGCGAAGAGTATCCAGTGACCCAAAACAAAGACATGCTGGTCCAGCTCAT 480
Db 511 AAATGTGTTGCGAAGAGTATCCAGTGACCCAAAACAAAGACATGCTGGTCCAGCTCAT 570
QY 481 GACGACTTCAAGATTTGAGGATGAATGGGATACATGCTGATGCTTTCGAAGTACTT 540
Db 571 GACGACTTCAAGATTTGAGGATGAATGGGATACATGCTGATGCTTTCGAAGTACTT 630
QY 541 GGCACCATCTCTCAAGTGGATCATCAATCCAACTATCAAGGCTCCAGTACGTTGT 600
Db 631 GGCACCATCTCTCAAGTGGATCATCAATCCAACTATCAAGGCTCCAGTACGTTGT 690
QY 601 GTGAAGAGTATCATTCGACAGGTCTCTTCAAGGTTAGATTACTTACAGATGAAGTCAAG 660
Db 691 GTGAAGAGTATCATTCGACAGGTCTCTTCAAGGTTAGATTACTTACAGATGAAGTCAAG 750
QY 661 ATCATTCATCTGACATAAAGCGGAAATATCTTGATGTGTGGATGATGCATATGTG 720
Db 751 ATCATTCATCTGACATAAAGCGGAAATATCTTGATGTGTGGATGATGCATATGTG 810
QY 721 AGAAGAAATGGCAGCTCAGGCCACTGAGTGGCAGAAAGAGGTGCTCTCTCTCAGGG 780
Db 811 AGAAGAAATGGCAGCTCAGGCCACTGAGTGGCAGAAAGAGGTGCTCTCTCTCAGGG 870
QY 781 TCTGAGTAGTACGGCTCCACAGCAGAAACCTATAGGAAATAATCTTAAAAACAAAAG 840
Db 871 TCTGAGTAGTACGGCTCCACAGCAGAAACCTATAGGAAATAATCTTAAAAACAAAAG 930
QY 841 AAAAACTGAAAGAAACAGAGAGCGCAGGTGAGTTATGGAGAGCGCTGCAGGAG 900
Db 931 AAAAACTGAAAGAAACAGAGAGCGCAGGTGAGTTATGGAGAGCGCTGCAGGAG 990
QY 901 ATAGAGAAATGGAGCGAGAGCTGAAAGGAAATAATAGAGAAACATCATCCTCAGCT 960
Db 991 ATAGAGAAATGGAGCGAGAGCTGAAAGGAAATAATAGAGAAACATCATCCTCAGCT 1050
QY 961 GCACCTTCCATGACAGGATGGCGAATATCTGCCAGAGGTGAAACTTAAACACACAGGA 1020
Db 1051 GCACCTTCCATGACAGGATGGCGAATATCTGCCAGAGGTGAAACTTAAACACACAGGA 1110
QY 1021 TTAGAGAGCGGCTGAGGCGAGAGCTGCAAGAGCAATGTTGAAGCTGAGGACACAGGAA 1080
Db 1111 TTAGAGAGCGGCTGAGGCGAGAGCTGCAAGAGCAATGTTGAAGCTGAGGACACAGGAA 1170
QY 1081 GAGAAAGAGATGCTGAGAAAGAAAACATTGAAAAAGATGAAGATGATGTAGATCAGGAA 1140
Db 1171 GAGAAAGAGATGCTGAGAAAGAAAACATTGAAAAAGATGAAGATGATGTAGATCAGGAA 1230
QY 1141 CTTGCGAATAGACCCCTAGTGGATGATATCCTTAAACCAATGCCATTTAGAAAT 1200
Db 1231 CTTGCGAATAGACCCCTAGTGGATGATATCCTTAAACCAATGCCATTTAGAAAT 1290
QY 1201 GGCCCTTCTCCTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAGACTGCCCA 1260
Db 1291 GGCCCTTCTCCTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAGACTGCCCA 1350
QY 1261 AATCTCTGAGGAATATATCTTGTATGAGCCCAATTCAGAAAGTGAATACATATAGCAGC 1320
Db 1351 AATCTCTGAGGAATATATCTTGTATGAGCCCAATTCAGAAAGTGAATACATATAGCAGC 1410
QY 1321 TCCTATGAACAAATTCATGTTGATTTGCCAAATGGACGACATAAATTTCCCGAGTCACAG 1380

Db 1411 TCCTATGAACAAATTCATGTTGAAATTCGCAAAATGGAGCATAAAATTTCCCGAGTCACAG 1470
QY 1381 TTCCACAGAGTTTTCACCTCGTTGTTCTCTGGATCCCTTTAGAACCTGTGGCTCTCGGCTCT 1440
Db 1471 TTCCACAGAGTTTTCACCTCGTTGTTCTCTGGATCCCTTTAGAACCTGTGGCTCTCGGCTCT 1530
QY 1441 GTGCTTTCTGAGGATCACCACTTACTGAGCAAGAGAGAGAGTCCATCCCATGACAGA 1500
Db 1531 GTGCTTTCTGAGGATCACCACTTACTGAGCAAGAGAGAGAGTCCATCCCATGACAGA 1590
QY 1501 AGCAGAGAGTTTTCAGCTCCAGTACCTGGGATTTGCCAAAAGCAAAAACCCCGGCGAGCT 1560
Db 1591 AGCAGAGAGTTTTCAGCTCCAGTACCTGGGATTTGCCAAAAGCAAAAACCCCGGCGAGCT 1650
QY 1561 GACTTGTGTGTAATCCCTGGATCCCGGAAATGCAGATAAAATTTAGAGTAAAAATTTGCT 1620
Db 1651 GACTTGTGTGTAATCCCTGGATCCCGGAAATGCAGATAAAATTTAGAGTAAAAATTTGCT 1710
QY 1621 GACTTGTGTAATGCTTGTGTTGTCATAAACACTTTCAGGAGAGATCCAGACGCTCAG 1680
Db 1711 GACTTGTGTAATGCTTGTGTTGTCATAAACACTTTCAGGAGAGATCCAGACGCTCAG 1770
QY 1681 TACCGCTCCATAGAGGTTTAAATAGGAGCGGGGTACAGACCCCTGGGACATCTGGAGC 1740
Db 1771 TACCGCTCCATAGAGGTTTAAATAGGAGCGGGGTACAGACCCCTGGGACATCTGGAGC 1830
QY 1741 ACGCGTGTATGTCATTTGAGCTGGCAACCGGAGATTAATTTGTTTGAACCACTTCTGGG 1800
Db 1831 ACGCGTGTATGTCATTTGAGCTGGCAACCGGAGATTAATTTGTTTGAACCACTTCTGGG 1890
QY 1801 GAAGACTATTCAGAGAGAGAGACACATAGCCACCATATAGAGCTGCTAGGAGATTT 1860
Db 1891 GAAGACTATTCAGAGAGAGAGACACATAGCCACCATATAGAGCTGCTAGGAGATTT 1950
QY 1861 CCAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTTCAATCGCAGAGGAGAA 1920
Db 1951 CCAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTTCAATCGCAGAGGAGAA 2010
QY 1921 CTGCGACACATCACAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTGTGGAAGATTT 1980
Db 2011 CTGCGACACATCACAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTGTGGAAGATTT 2070
QY 1981 GGCTGGCCCCATGAAGATGCTGCACAGTTTACAGATTTCCCTGATCCCGATGTTAGAAATG 2040
Db 2071 GGCTGGCCCCATGAAGATGCTGCACAGTTTACAGATTTCCCTGATCCCGATGTTAGAAATG 2130
QY 2041 GTTCCAGAAAAACAGAGCCTCAGCTGGCAATGCCCTTCGGCATCTTGGTTGAATTTCT 2097
Db 2131 GTTCCAGAAAAACAGAGCCTCAGCTGGCAATGCCCTTCGGCATCTTGGTTGAATTTCT 2187

RESULT 2

AX666205

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1..3253

/organism="Homo sapiens"

AX666205

Sequence 1 from Patent WO02057458.

AX666205.1 GI:29291001

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Yan, C., Abu-Threideh, J., Shao, W., Merklov, G., di Francesco, V. and

Beasley, E.M.

Isolated human kinase proteins, nucleic acid molecules encoding

human kinase proteins, and uses thereof

Patent: WO 02057458-A, 1 25-JUL-2002;

PE Corporation (NY) (US)

Location/Qualifiers

1..3253

/organism="Homo sapiens"

AX666205

Sequence 1 from Patent WO02057458.

AX666205.1 GI:29291001

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Yan, C., Abu-Threideh, J., Shao, W., Merklov, G., di Francesco, V. and

Beasley, E.M.

Isolated human kinase proteins, nucleic acid molecules encoding

human kinase proteins, and uses thereof

Patent: WO 02057458-A, 1 25-JUL-2002;

PE Corporation (NY) (US)

Location/Qualifiers

1..3253

/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 2097; DB 6; Length 3253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2097; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAGCTCCGGAAAGTGTGCGCCATTTCAGGCCGCGAAAGCGGAGCGCGAAAGAGAGAAA 60
DB |||
QY 91 ATGAGCTCCCGAAAGTGTGCGCCATTTCAGGCCGCGAAAGCGGAGCGCGAAAGAGAGAAA 150
DB |||
QY 61 CATCCGAAAAAGCCGAGGCTCAACAGAAAGTCTCTTTAGTTCTCTCTCTCTCTCTCTCTCTCT 120
DB |||
QY 151 CATCCGAAAAAGCCGAGGCTCAACAGAAAGTCTCTTTAGTTCTCTCTCTCTCTCTCTCTCTCT 210
DB |||
QY 121 CCACCAACCAACCCGCGACCTTTGCGAGACCCCGACACCCCGGAGCCGAGGAGGAGATC 180
DB |||
QY 211 CCACCAACCAACCCGCGACCTTTGCGAGACCCCGACACCCCGGAGCCGAGGAGGAGATC 270
DB |||
QY 181 CTGGGATCAGATGATGAGGAGCAAGAGACCTTGCAGACTACTGCGAAGGTGATATCAT 240
DB |||
QY 271 CTGGGATCAGATGATGAGGAGCAAGAGACCTTGCAGACTACTGCGAAGGTGATATCAT 330
DB |||
QY 241 CCAGTGAAATTTGGAGACCTTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATCG 300
DB |||
QY 331 CCAGTGAAATTTGGAGACCTTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATCG 390
DB |||
QY 301 GGGCACTTCTCTACTGTCTGGCTGTGCTGGGATATGCGAGGGGAAAGATTTGTTGCAATG 360
DB |||
QY 391 GGGCACTTCTCTACTGTCTGGCTGTGCTGGGATATGCGAGGGGAAAGATTTGTTGCAATG 450
DB |||
QY 361 AAAGTTGTAAGAGTCCGAGACATTTACGGAGACAGCTTGGATGGAATGAAATTAATTTGCTC 420
DB |||
QY 451 AAAGTTGTAAGAGTCCGAGACATTTACGGAGACAGCTTGGATGGAATGAAATTAATTTGCTC 510
DB |||
QY 421 AAATGTGTTTCGAAAGTGCATCCAGTGACCCAAACAAAGACATGTTGCTCCAGCTCATTT 480
DB |||
QY 511 AAATGTGTTTCGAAAGTGCATCCAGTGACCCAAACAAAGACATGTTGCTCCAGCTCATTT 570
DB |||
QY 481 GAGCACTTCAAGATTTTCAGGCATGAATGGGATACATGTTCTGATGTTCTTTCGAAGTACTTT 540
DB |||
QY 571 GAGCACTTCAAGATTTTCAGGCATGAATGGGATACATGTTCTGATGTTCTTTCGAAGTACTTT 630
DB |||
QY 541 GGCACCATCTCTCTAAGTGGATCATCAATCCAACTATCAAGGCCCTCCAGTAGCTTTGT 600
DB |||
QY 631 GGCACCATCTCTCTAAGTGGATCATCAATCCAACTATCAAGGCCCTCCAGTAGCTTTGT 690
DB |||
QY 601 GTGAAGAGTATCATTTGACAGAGTCTTCAAGGTTAGATTCTTACACAGTAAGTGCAG 660
DB |||
QY 691 GTGAAGAGTATCATTTGACAGAGTCTTCAAGGTTAGATTCTTACACAGTAAGTGCAG 750
DB |||
QY 661 ATCATTTCATCTACATAAAGCGGAAAAATATCTTGATGTGTGATGATGATCATATGTG 720
DB |||
QY 751 ATCATTTCATCTACATAAAGCGGAAAAATATCTTGATGTGTGATGATGATCATATGTG 810
DB |||
QY 721 AGAAGAAATGGAGCTGAGGCGCATGAGTGGCGAAGAGAGGTGCTCTCTCTCTTCAGGG 780
DB |||
QY 811 AGAAGAAATGGAGCTGAGGCGCATGAGTGGCGAAGAGAGGTGCTCTCTCTCTTCAGGG 870
DB |||
QY 781 TCTGCAAGTACAGTACGCTCCAGCAGAGACCTTATAGGAAATATCTTAAACCAAAAG 840
DB |||
QY 871 TCTGCAAGTACAGTACGCTCCAGCAGAGACCTTATAGGAAATATCTTAAACCAAAAG 930
DB |||
QY 841 AAAAAACTGAAAAAGAAACAGAGAGCGAGGTGAGTTATTGGAGAGCGCTTCGAGGAG 900
DB |||
QY 931 AAAAACTGAAAAAGAAACAGAGAGCGAGGTGAGTTATTGGAGAGCGCTTCGAGGAG 990
DB |||
QY 901 ATAGAAGAAATGGAGCGAGAGCTGAAGGAAATTAATAGAGAAACATCACTCAGCT 960
DB |||
QY 991 ATAGAAGAAATGGAGCGAGAGCTGAAGGAAATTAATAGAGAAACATCACTCAGCT 1050
DB |||
QY 961 GCACCTTCCAAATGACCAAGATGGCGAATATCTGCCAGAGGTGAAACTTAAACCAACAGGA 1020
```

```
DB |||
QY 1051 GCACCTTCCAAATGACCAAGATGGCGAATATCTGCCAGAGGTGAAACTTAAACCAACAGGA 1110
DB |||
QY 1021 TTAGAGGAGCGCGCTGAGCGAGAGACTGCAAAAGGACAAATGGTGAAGCTGAGGACCAAGAA 1080
DB |||
QY 1111 TTAGAGGAGCGCGCTGAGCGAGAGACTGCAAAAGGACAAATGGTGAAGCTGAGGACCAAGAA 1170
DB |||
QY 1081 GAGAAAGAGATGCTGAGAAAGAAACATTTGAAAAGATGAAAGATGATGTAGATCAGGAA 1140
DB |||
QY 1171 GAGAAAGAGATGCTGAGAAAGAAACATTTGAAAAGATGAAAGATGATGTAGATCAGGAA 1230
DB |||
QY 1141 CTTTGGAAACATAGACCTTACGTGGATAGAAATCACTTAAACCAATGGCCATATTGAGAAAT 1200
DB |||
QY 1231 CTTTGGAAACATAGACCTTACGTGGATAGAAATCACTTAAACCAATGGCCATATTGAGAAAT 1290
DB |||
QY 1201 GGGCCATTTCTACTGGAGCAGCAACTGGACGATGATGATGAAAGAGACTGCCCA 1260
DB |||
QY 1291 GGGCCATTTCTACTGGAGCAGCAACTGGACGATGATGATGAAAGAGACTGCCCA 1350
DB |||
QY 1261 AATCTGAGGAATATAATCTTGTGATGAGCCAAATGCAAGAAAGTATTTACACATATAGCAGC 1320
DB |||
QY 1351 AATCTGAGGAATATAATCTTGTGATGAGCCAAATGCAAGAAAGTATTTACACATATAGCAGC 1410
DB |||
QY 1321 TCCTATGAACAAATCAATGGTGAATTTGCCAAATGGACGACATATAAAATTTCCGAGTCACAG 1380
DB |||
QY 1411 TCCTATGAACAAATCAATGGTGAATTTGCCAAATGGACGACATATAAAATTTCCGAGTCACAG 1470
DB |||
QY 1381 TTCCAGAGATTTTCCAACCTGCTGTTCTCTGGATCCTTAGAACCTGTGGCTGCGGCTCT 1440
DB |||
QY 1471 TTCCAGAGATTTTCCAACCTGCTGTTCTCTGGATCCTTAGAACCTGTGGCTGCGGCTCT 1530
DB |||
QY 1441 GTGCTTTCTGAGGATTCACCTTACTGAGCAAGAGGAGAGAGAGTCCATCCCATGACAGA 1500
DB |||
QY 1531 GTGCTTTCTGAGGATTCACCTTACTGAGCAAGAGGAGAGAGAGTCCATCCCATGACAGA 1590
DB |||
QY 1501 AGCAGAACGGTTTTCAGCTTCCAGTACTGGGATTTGCCAAAAGCAAAAACCCGGGAGCT 1560
DB |||
QY 1591 AGCAGAACGGTTTTCAGCTTCCAGTACTGGGATTTGCCAAAAGCAAAAACCCGGGAGCT 1650
DB |||
QY 1561 GACTTGTGTTGGTGAATCCCTGGATCCCGGAAATGACAGATAAAATTTAGAGTAAAAATTTGCT 1620
DB |||
QY 1651 GACTTGTGTTGGTGAATCCCTGGATCCCGGAAATGACAGATAAAATTTAGAGTAAAAATTTGCT 1710
DB |||
QY 1621 GACTTGGGAAATGCTTGTGGTGCATATAAACAATTCA CGGAAGACATCCAGAGCGCTCAG 1680
DB |||
QY 1711 GACTTGGGAAATGCTTGTGGTGCATATAAACAATTCA CGGAAGACATCCAGAGCGCTCAG 1770
DB |||
QY 1681 TACCGCTCCATAGAGGTTTAAATAGGAGCGGGTACAGCACCCCTGCGGACATCTCGAGC 1740
DB |||
QY 1771 TACCGCTCCATAGAGGTTTAAATAGGAGCGGGTACAGCACCCCTGCGGACATCTCGAGC 1830
DB |||
QY 1741 ACGGGCTGTATGGCATTTGAGCTGGCAACGGGAGATTTATTTGTTGAAACCAATTTCTGGG 1800
DB |||
QY 1831 ACGGGCTGTATGGCATTTGAGCTGGCAACGGGAGATTTATTTGTTGAAACCAATTTCTGGG 1890
DB |||
QY 1801 GAAGACTATTTCAGAGAGCAAGACACATAGCCCA CATATAGAGCTGTAGCAGTATT 1860
DB |||
QY 1891 GAAGACTATTTCAGAGAGCAAGACACATAGCCCA CATATAGAGCTGTAGCAGTATT 1950
DB |||
QY 1861 CCAAGGCACTTTGCTTATCTCGAAAAATTTCTCGGAAATTTCTCAATCGCAGAGAGAA 1920
DB |||
QY 1951 CCAAGGCACTTTGCTTATCTCGAAAAATTTCTCGGAAATTTCTCAATCGCAGAGAGAA 2010
DB |||
QY 1921 CTGCGACACATCAACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTATCTTGTGAAAAAGTAT 1980
DB |||
QY 2011 CTGCGACACATCAACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTGTGAAAAAGTAT 2070
DB |||
QY 1981 GCGTGGCCCCATGAAGATGCTGCA CAGTTTACAGATTTCTGATCCCGATGTTAGAAATG 2040
DB |||
QY 2071 GCGTGGCCCCATGAAGATGCTGCA CAGTTTACAGATTTCTGATCCCGATGTTAGAAATG 2130
DB |||
QY 2041 GTTCAGAAAAACGAGCCTCAGCTGGGAAATGCGCTTCGGCATCTCTGGTTGATTTCT 2097
DB |||
```


DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 167 Row: 1 Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507220.

FEATURES

Source

Location/Qualifiers
1..2577
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:87435 IMAGE:5265599"
/tissue_type="Testis"
/clone_lib="NIH MGC 97"
/lab_host="DH10B"
/note="Vector: pBluescript"

gene

1..2577
/gene="SRPK2"
/note="synonym: SPRSK2"
/db_xref="LocusID:6733"
/db_xref="MIM:602980"
182..2248
/gene="SRPK2"
/codon_start=1
/product="SRPK2 protein"
/protein_id="AAH68547.1"
/db_xref="GI:46250445"
/db_xref="LocusID:6733"
/db_xref="MIM:602980"

CDS

/translation="MSVNSEKSSSERPEPOQKAPLPPPPPPPPPPPPPPDPPTPEE
PEETLGSDDEQEDPADYCKGYHPVKIGDLFNRYHVRKLGWGFSTVWLCWDMQ
GKRFVAMKVGAQHWITLALDEIKLLKCRSDRSDPNKMWOLIDDFKISGMNGI
HVMVFVLGHLLHWI1KSNYQGLPVRCKSI1RQVLQGLDYLHSCKKI1HTDIKPE
N1LMCVDDAYVRMAEATEWQKAGAPPSAVSTAPOQKPIGKISNKKKLLKKQ
KQAELEKRLQIELEAREAKI1EENITSAAFSNDDQDEYCEPVKLTITGLEEAA
EAATKAGAEDEQKEADEKENIEKDDVDVLANIDPTWIESPKTHIENGPF
SLEQLDDEDDDECPNPEYNDLDPNAESDYTSSYEQFNGLPNRHKIPESQF
PFFSLSFGSELPVACSVLSEGSPLTEQESSPSHDSRTVSASSGDLPKAKTRA
ADLLNPIDNRADIRVKIADLGNACVWHKFTED1OTROYRSLEVLIGAGYSTPAD
INSTACMAFELATDYLPEPHSGEDYSRDEDHAI1ELLGSI1PHEFALSQKYSREFF
NRRBLRHITKLKPLGLFDVLVEKYGWPHEDAAQFTDFLIPMLEMWPEKRASGECLE
HPWLNS"
419..847
/gene="SRPK2"
/note="S TK; Region: Serine/Threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"
1757..2239
/gene="SRPK2"
/note="S TK; Region: Serine/Threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"

misc_feature

419..847
/gene="SRPK2"
/note="S TK; Region: Serine/Threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"
1757..2239
/gene="SRPK2"
/note="S TK; Region: Serine/Threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"

misc_feature

1757..2239
/gene="SRPK2"
/note="S TK; Region: Serine/Threonine protein kinases,
catalytic domain. Phosphotransferases of the serine or
threonine-specific kinase subfamily. The enzymatic
activity of these protein kinases is controlled by
phosphorylation of specific residues in the activation
segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"

ORIGIN

Query Match

96.5%; Score 2023.8; DB 9; Length 2577;

	Best Local Similarity	99.7%; Pred. No. 0;	Mismatches	0;	Indels	0;	Gaps	0;
	Matches	2028;	Conservative	0;				
QY	63	TCGAGAAAGCGCGGAGCCTCAACAGAAAGCTCTTTAGTTCCTCTCCACCGCCACC	122					
DB	211	TTGAGAAAGCGCGGAGCCTCAACAGAAAGCTCTTTAGTTCCTCTCCACCGCCACC	270					
QY	123	ACCACACACACCGCCACCTTTTGCAGAGCCCAACACCCCGGAGCCAGAGGAGATCCT	182					
DB	271	ACCACACACACCGCCACCTTTTGCAGAGCCCAACACCCCGGAGCCAGAGGAGATCCT	330					
QY	183	GGGATCAGATGATGAGGAGCAAGAGACCTTCGCGACTCTCTCAAGGTGGATATCATCC	242					
DB	331	GGGATCAGATGATGAGGAGCAAGAGACCTTCGCGACTCTCTCAAGGTGGATATCATCC	390					
QY	243	AGTGAATAATGGAGACCTCTCAATGGCGCGTATCATGTTATAGAAAGCTTGGATGGG	302					
DB	391	AGTGAATAATGGAGACCTCTCAATGGCGCGTATCATGTTATAGAAAGCTTGGATGGG	450					
QY	303	GCACCTTCTACTGTCTGTCTGTGGATATGAGGGGAAAGATTTGTTGCAATGAA	362					
DB	451	GCACCTTCTACTGTCTGTCTGTGGATATGAGGGGAAAGATTTGTTGCAATGAA	510					
QY	363	AGTTCTAAAGTGCCTCAGACATTTATCGGAGACAGCTTTGGATGAAATAAATTTGCTCA	422					
DB	511	AGTTCTAAAGTGCCTCAGACATTTATCGGAGACAGCTTTGGATGAAATAAATTTGCTCA	570					
QY	423	ATGCTCTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGTTGTTGCAATGCA	482					
DB	571	ATGCTCTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGTTGTTGCAATGCA	630					
QY	483	CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGTCATGCTTTCGAAAGTACTTGG	542					
DB	631	CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGTCATGCTTTCGAAAGTACTTGG	690					
QY	543	CCACATCTCTCAAGTGGATCATCAAAATCCAATATCAAGGCCTCCAGTAGTCTGTGTGT	602					
DB	691	CCACATCTCTCAAGTGGATCATCAAAATCCAATATCAAGGCCTCCAGTAGTCTGTGTGT	750					
QY	603	GAAGAGTATCATTCACAGAGTCTCTCAAGGTTTACATTTACACAGTAACTGCAAGAT	662					
DB	751	GAAGAGTATCATTCACAGAGTCTCTCAAGGTTTACATTTACACAGTAACTGCAAGAT	810					
QY	663	CATTTCATCTGCATATAAAGCGGAAATATCTTGTATGTTGTTGATGATGCATATGTGAG	722					
DB	811	CATTTCATCTGCATATAAAGCGGAAATATCTTGTATGTTGTTGATGATGCATATGTGAG	870					
QY	723	AAGAAATGGCAGTGAAGGCACTGAGTGGAGAAAGAGGTGCTCTCTCTCTTCAGGGTTC	782					
DB	871	AAGAAATGGCAGTGAAGGCACTGAGTGGAGAAAGAGGTGCTCTCTCTCTTCAGGGTTC	930					
QY	783	TGCAGTGAAGTACGGCTCCACAGCAAAACCTATAGGAAATAATCTTAAACACAAAGAA	842					
DB	931	TGCAGTGAAGTACGGCTCCACAGCAAAACCTATAGGAAATAATCTTAAACACAAAGAA	990					
QY	843	AAAACCTGAAAAGAAACAGAGAGCAGGCTGAGTATTATTGGAGAGCGCTCAGGAGAT	902					
DB	991	AAAACCTGAAAAGAAACAGAGAGCAGGCTGAGTATTATTGGAGAGCGCTCAGGAGAT	1050					
QY	903	AGAGAAATTTGGAGCGAGAGCTGAAAGGAAATAATAGAGAAACATCATCCTCAGCTGC	962					
DB	1051	AGAGAAATTTGGAGCGAGAGCTGAAAGGAAATAATAGAGAAACATCATCCTCAGCTGC	1110					
QY	963	ACCTTCCAAATGACAGGATGGCGAATCTGCCAGAGGTGAAATCTTAAACACAAAGATTT	1022					
DB	1111	ACCTTCCAAATGACAGGATGGCGAATCTGCCAGAGGTGAAATCTTAAACACAAAGATTT	1170					
QY	1023	AGAGAGGCGGCTGAGGCGAGACCTGCAAGAGCAATGTTGAAGCTGAGGACAGGAGAGA	1082					
DB	1171	AGAGAGGCGGCTGAGGCGAGACCTGCAAGAGCAATGTTGAAGCTGAGGACAGGAGAGA	1230					
QY	1083	GAAAGAGATGCTGAGAAAGAAAACATTTGAAAAGATGATGATGATGATGATGATGATGAT	1142					

Db	1231	GAAGAAGATGCTGAGAAAGAAAAACATTGAAAAAGATGAAGATGATGTAGATCAGGAAC	1290
Qy	1143	TGCGACATAGACCCCTACCTGATAGATAAATCACTAAACCAATGGCCATATTGAGATCG	1202
Db	1291	TGCGAACATAGACCCCTACCTGATAGATAAATCACTAAACCAATGGCCATATTGAGATCG	1350
Qy	1203	CCCATCTCAGCTGGAGCAGCAACTGACGATGAAGATGATGATGAAGAGACTGCCCAA	1262
Db	1351	CCCATCTCAGCTGGAGCAGCAACTGACGATGAAGATGATGATGAAGAGACTGCCCAA	1410
Qy	1263	TCTGAGGAATATAATCTTTGATGAGCAAAATGAGAAAGTATTACATATAGCAGCTC	1322
Db	1411	TCTGAGGAATATAATCTTTGATGAGCAAAATGAGAAAGTATTACATATAGCAGCTC	1470
Qy	1323	CTATGAACATTCNAATGGTGAATTCGCAATGGACGACATAAATCCCGAGTCACAGTT	1382
Db	1471	CTATGAACATTCNAATGGTGAATTCGCAATGGACGACATAAATCCCGAGTCACAGTT	1530
Qy	1383	CCGAGATTTTCCACCTCTGTCTCTGGATCCTTAGAACCTGTGGCTCGGGCTCTCT	1442
Db	1531	CCGAGATTTTCCACCTCTGTCTCTGGATCCTTAGAACCTGTGGCTCGGGCTCTCT	1590
Qy	1443	GCTTCTGAGGATCACCTTACTGAGCAAGAGGAGCAGTCCATCCCATGACAGAAG	1502
Db	1591	GCTTCTGAGGATCACCTTACTGAGCAAGAGGAGCAGTCCATCCCATGACAGAAG	1650
Qy	1503	CAGAACGTTTTCAGCTCAGTACTGGGGATTTGCCAAAGCAAAACCCGGGCGACTGA	1562
Db	1651	CAGAACGTTTTCAGCTCAGTACTGGGGATTTGCCAAAGCAAAACCCGGGCGACTGA	1710
Qy	1563	CTTGTGGTGAATCCCTCGATCCGGGAATGCAGATAAATTAGAGTAAATAATGCTGA	1622
Db	1711	CTTGTGGTGAATCCCTCGATCCGGGAATGCAGATAAATTAGAGTAAATAATGCTGA	1770
Qy	1623	CCTGGGAAATGCTTGTGGTGCATAAACACTTTCAGGAAGACATCCAGACGGTCACTA	1682
Db	1771	CCTGGGAAATGCTTGTGGTGCATAAACACTTTCAGGAAGACATCCAGACGGTCACTA	1830
Qy	1683	CGCTCCATAGAGTTTAAATAGAGCGGGGTACAGCACCCCTCGGGAATCTGAGAGCAC	1742
Db	1831	CGCTCCATAGAGTTTAAATAGAGCGGGGTACAGCACCCCTCGGGAATCTGAGAGCAC	1890
Qy	1743	GCCTGTATGGCATTTGACCTGCACCGGAGATTTATTTGTTGAACCAATCTGGGGA	1802
Db	1891	GCCTGTATGGCATTTGACCTGCACCGGAGATTTATTTGTTGAACCAATCTGGGGA	1950
Qy	1803	AGACTATTCAGAGACGAAGACACATAGCCACACATCATAGAGCTGCTAGGCAATATCC	1862
Db	1951	AGACTATTCAGAGACGAAGACACATAGCCACACATCATAGAGCTGCTAGGCAATATCC	2010
Qy	1863	AAGGCACTTTGCTATCTGGAATAATTTCTGGGAATTTCTCAATCGCAGAGGAAC	1922
Db	2011	AAGGCACTTTGCTATCTGGAATAATTTCTGGGAATTTCTCAATCGCAGAGGAAC	2070
Qy	1923	GGGACACATCACAGCTCAAGCCCTGGAGCCCTTTGATGATCTGTTGGAAGTATGG	1982
Db	2071	GGGACACATCACAGCTCAAGCCCTGGAGCCCTTTGATGATCTGTTGGAAGTATGG	2130
Qy	1983	CTGGCCCCATGAAGATGCTGCACAGTTTACAGATTTCTTGATCCCGATGTTAGAAATGG	2042
Db	2131	CTGGCCCCATGAAGATGCTGCACAGTTTACAGATTTCTTGATCCCGATGTTAGAAATGG	2190
Qy	2043	TCAGAAAAACGAGCTCAGCTGGGAAATGCCCTTCGGCATCTCTGGTGAATTTCT	2097
Db	2191	TCAGAAAAACGAGCTCAGCTGGGAAATGCCCTTCGGCATCTCTGGTGAATTTCT	2245

RESULT 5
AR077238
LOCUS
DEFINITION
ACCESSION
VERSION

Sequence 8 from patent US 5962232.
AR077238
AR077238
AR077238.1

2791 bp
DNA
linear
PAT 31-AUG-2000

GI:10003984

KEYWORDS	Unknown.									
SOURCE	Unknown.									
ORGANISM	Unclassified.									
REFERENCE	1 (bases 1 to 2791)									
AUTHORS	Bandman,O., Hillman,J.L., Lal,P., Akerblom,I.E., Shah,P., Corley,N.C. and Guegler,K.J.									
TITLE	Protein kinase molecules									
JOURNAL	Patent: US 5962232-A 8 05-OCT-1999;									
FEATURES	Location/Qualifiers									
source	1..2791									
	/organism="unknown"									
	/mol_type="unassigned DNA"									
ORIGIN										
Query Match	96.5%; Score 2023.8; DB 6; Length 2791;									
Best local Similarity	99.7%; Pred. No. 0;									
Matches 2028; Conservative	0; Mismatches 7; Indels 0; Gaps 0;									
Qy	63	TCGAAAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCCTCCTCCACCGCCACC	122							
Db	202	TTCAGAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCCTCCTCCACCGCCACC	261							
Qy	123	ACCACACACCCGACCTTTGCCAGACCCACACCCCGGAGCCAGAGAGAGATCCT	182							
Db	262	ACCACACACCCGCCCTTTTGCAGACCCACACCCCGGAGCCAGAGAGAGATCCT	321							
Qy	183	GGGATCAGATGATGAGGACCAAGAGACCTCTCGGACTACTGCAAAAGGTGGATATCATCC	242							
Db	322	GGGATCAGATGATGAGGACCAAGAGACCTCTCGGACTACTGCAAAAGGTGGATATCATCC	381							
Qy	243	AGTGAATTTGGAGACCTCTTCAATGCCCGTATCATGTTATTAGAAAGCTTGGATGGGG	302							
Db	382	AGTGAATTTGGAGACCTCTTCAATGCCCGTATCATGTTATTAGAAAGCTTGGATGGGG	441							
Qy	303	GCACCTTCTACTGCTGTGGGATATGCAAGGGGAAAAGATTTGTTGCAATGAA	362							
Db	442	GCACCTTCTACTGCTGTGGGATATGCAAGGGGAAAAGATTTGTTGCAATGAA	501							
Qy	363	AGTTGTAAAAAGTGCCAGCATTATACGGAGACAGCCTTTGGATGAAATAAAATTTGCTCAA	422							
Db	502	AGTTGTAAAAAGTGCCAGCATTATACGGAGACAGCCTTTGGATGAAATAAAATTTGCTCAA	561							
Qy	423	ATGTTGTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA	482							
Db	562	ATGTTGTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA	621							
Qy	483	CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGTCATGGTCTTTCGAAAGTACTTGG	542							
Db	622	CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGTCATGGTCTTTCGAAAGTACTTGG	681							
Qy	543	CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCCTCCAGTAGCTTTGTGT	602							
Db	682	CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCCTCCAGTAGCTTTGTGT	741							
Qy	603	GAGAGTATCATTCGACAGGTCTCTCAAGGGTTAGATTTACTTACACAGTAAGTGCAGAT	662							
Db	742	GAGAGTATCATTCGACAGGTCTCTCAAGGGTTAGATTTACTTACACAGTAAGTGCAGAT	801							
Qy	663	CATTTCATCTGACATAAAGCCGAAAATATCTTTGATGTGTGTGGATGATGCATATGTGAG	722							
Db	802	CATTTCATCTGACATAAAGCCGAAAATATCTTTGATGTGTGTGGATGATGCATATGTGAG	861							
Qy	723	AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCAGGGTC	782							
Db	862	AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCAGGGTC	921							
Qy	783	TGCAGTATGAGTACGGCTCCACAGCAGAAAACTATAGGAAAAATATCTTAAAAACAAGAA	842							
Db	922	TGCAGTATGAGTACGGCTCCACAGCAGAAAACTATAGGAAAAATATCTTAAAAACAAGAA	981							
Qy	843	AAAACTGAAAAAAGAACAGAGAGCGGCTGAGTTATTGGAGAGCGCCTTCAGGAGAT	902							

Db 442 GCACCTTCTCTGCTGCTGGGATATGACAGGGGAAAAAGATTTGTTGCAATGAA 501
Qy 363 AGTTGTAAGAGTCCAGCATATACGAGACAGCCTTGGATGAATAAATTTGCTCAA 422
Db 502 AGTTGTAAGAGTCCAGCATATACGAGACAGCCTTGGATGAATAAATTTGCTCAA 561
Qy 423 ATGTGTTTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATGCA 482
Db 562 ATGTGTTTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATGCA 621
Qy 483 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCATGGTCTTCGAAAGTACTTGG 542
Db 622 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCATGGTCTTCGAAAGTACTTGG 681
Qy 543 CCACCATCTCCTCAAGTGTGATCATCAATCCAACTATCAAGGCCTCCAGTACGTTGTGT 602
Db 682 CCACCATCTCCTCAAGTGTGATCATCAATCCAACTATCAAGGCCTCCAGTACGTTGTGT 741
Qy 603 GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTACTTTACACAGTAAGTGAAGAT 662
Db 742 GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTACTTTACACAGTAAGTGAAGAT 801
Qy 663 CATTCATCTGACATAAAGCCGGAATAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 722
Db 802 CATTCATCTGACATAAAGCCGGAATAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 861
Qy 723 AAGATGGCAGTGTAGGCGACATGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCTCTCTCTCT 782
Db 862 AAGATGGCAGTGTAGGCGACATGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCTCTCTCTCT 921
Qy 783 TGCAGTGTAGTACGGCTCCACACGAGAAACCTATAGGAAATAATCTTAAACCAAAAGAA 842
Db 922 TGCAGTGTAGTACGGCTCCACACGAGAAACCTATAGGAAATAATCTTAAACCAAAAGAA 981
Qy 843 AAAACTGAAAAAGAAACAGAGCAGGCTGAGTTATTGGAGAGCGGCTCCAGGAGAT 902
Db 982 AAAACTGAAAAAGAAACAGAGCAGGCTGAGTTATTGGAGAGCGGCTCCAGGAGAT 1041
Qy 903 AGAAGATTTGGAGCGAGAGAGCTGAAAGGAAATAATAGAAAGAAACATCACCTCAGCTGC 962
Db 1042 AGAAGATTTGGAGCGAGAGCTGAAAGGAAATAATAGAAAGAAACATCACCTCAGCTGC 1101
Qy 963 ACCTTCCATGACAGGATGGCGAATACCTGCCAGAGGTGAACCTTAAACCAACAGGATT 1022
Db 1102 ACCTTCCATGACAGGATGGCGAATACCTGCCAGAGGTGAACCTTAAACCAACAGGATT 1161
Qy 1023 AGAGGAGCGGCTGAGGACAGACTGCAAGAGCAATGCTGAAGCTGAGGACCAAGAGA 1082
Db 1162 AGAGGAGCGGCTGAGGACAGACTGCAAGAGCAATGCTGAAGCTGAGGACCAAGAGA 1221
Qy 1083 GAAAGAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGATGATGATGATGAT 1142
Db 1222 GAAAGAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGATGATGATGATGAT 1281
Qy 1143 TCGAACATAGACCTTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1202
Db 1282 TCGAACATAGACCTTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1341
Qy 1203 CCATCTTCACTGAGAGCAGCACTGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1262
Db 1342 CCATCTTCACTGAGAGCAGCACTGCAAGATGATGATGATGATGATGATGATGATGATGATGAT 1401
Qy 1263 TCCTGAGGAATAATCTTGTATGAGCCAAATCCAGAAAGTGTATTACATATAGCAGCTC 1322
Db 1402 TCCTGAGGAATAATCTTGTATGAGCCAAATCCAGAAAGTGTATTACATATAGCAGCTC 1461
Qy 1323 CTATGAACATTTCAATGTGTGATTTGCCAAATGGACGACATAAATTTCCGAGTCAAGTT 1382
Db 1462 CTATGAACATTTCAATGTGTGATTTGCCAAATGGACGACATAAATTTCCGAGTCAAGTT 1521
Qy 1383 CCCAGATTTTCCACTGCTGCTTCTCTGATCTCTAGAACCTGTGGCTGGGCTCTGT 1442
Db 1522 CCCAGATTTTCCACTGCTGCTTCTCTGATCTCTAGAACCTGTGGCTGGGCTCTGT 1581

Qy 1443 GCTTTCTGAGGGATCACCACTTACTGAGCAAGAGGAGCAGTCCATCCATCCATCAGAGAAG 1502
Db 1582 GCTTTCTGAGGGATCACCACTTACTGAGCAAGAGGAGCAGTCCATCCATCCATCAGAGAAG 1641
Qy 1503 CAGAACGGTTTCAGCCTCCAGTACTGGGATTTGCCAAAGCAAAAAACCCGGGCGAGCTGA 1562
Db 1642 CAGAACGGTTTCAGCCTCCAGTACTGGGATTTGCCAAAGCAAAAAACCCGGGCGAGCTGA 1701
Qy 1563 CTTGTTGTGTAATCCCTCGGATCCGCGGAAATCCAGATAAAAAATTTAGAGTAAAAATTCCTGA 1622
Db 1702 CTTGTTGTGTAATCCCTCGGATCCGCGGAAATCCAGATAAAAAATTTAGAGTAAAAATTCCTGA 1761
Qy 1623 CCGTGGAAATGCTTGTGGTGCATAAACAACCTTTCAGGAGAGATCCAGACGGCTCAGTA 1682
Db 1762 CCGTGGAAATGCTTGTGGTGCATAAACAACCTTTCAGGAGAGATCCAGACGGCTCAGTA 1821
Qy 1683 CGCTCCCATAGAGTTTAAATAGGAGCGGGTACAGACCCCTCGCGACATCTGGAGCAC 1742
Db 1822 CGCTCCCATAGAGTTTAAATAGGAGCGGGTACAGACCCCTCGCGACATCTGGAGCAC 1881
Qy 1743 GCGGTGTATGCAATTTGAGCTGGCAACCGGAGATTTATTTGTTGAAACCAATTTCTGGGA 1802
Db 1882 GCGGTGTATGCAATTTGAGCTGGCAACCGGAGATTTATTTGTTGAAACCAATTTCTGGGA 1941
Qy 1803 AGACTATTTCCAGAGAGCAAGACCAATAGCCCAATCATAGAGCTGCTAGGCAAGTATTC 1862
Db 1942 AGACTATTTCCAGAGAGCAAGACCAATAGCCCAATCATAGAGCTGCTAGGCAAGTATTC 2001
Qy 1863 AAGGCACATTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTCAATCGCAGAGGAGAACT 1922
Db 2002 AAGGCACATTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTCAATCGCAGAGGAGAACT 2061
Qy 1923 GCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTAATTTGTGGAAGATATGG 1982
Db 2062 GCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTAATTTGTGGAAGATATGG 2121
Qy 1983 CTGGCCCATAGAAATGCTGCAAGTTTACAGATTTCTTGATCCCGATGTTAGAAATGGT 2042
Db 2122 CTGGCCCATAGAAATGCTGCAAGTTTACAGATTTCTTGATCCCGATGTTAGAAATGGT 2181
Qy 2043 TCCAGAAAAACGAGCCTCAGCTGCGGAATGCTTCCGCACTCTTGTGGAATTTCT 2097
Db 2182 TCCAGAAAAACGAGCCTCAGCTGCGGAATGCTTCCGCACTCTTGTGGAATTTCT 2236

RESULT 7

Q0715987
LOCUS Q0715987 3708 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 1921 from Patent WO02068579.
ACCESSION Q0715987
VERSION Q0715987.1 GI:42276844
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
AUTHORS Kites, such as nucleic acid arrays, comprising a majority of
TITLE humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 1921 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
source 1..3708
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 95.7%; Score 2006.6; DB 6; Length 3708;
Best Local Similarity 99.7%; Pred. No. 0;

Matches	2031:	Conservative	0;	Mismatches	4;	Indels	2;	Gaps	2;
Qy	63	TCGAAAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCTCCACCGCAACC	122						
Db	202	TTCAGAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCTCCACCGCAACC	261						
Qy	123	ACCACCAACCGCCACTTTGCCAGACCCCAACCCCGGAGCCAGAGGAGATCCT	182						
Db	262	ACCACCAACCGCCACTTTGCCAGACCCCAACCCCGGAGCCAGAGGAGATCCT	321						
Qy	183	GGATCAGATGATGAGGAGCAAGAGGACCTTCGGGACTATCTGCAAAAGGTGGATATCATCC	242						
Db	322	GGATCAGATGATGAGGAGCAAGAGGACCTTCGGGACTATCTGCAAAAGGTGGATATCATCC	381						
Qy	243	AGTGAAAAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGGG	302						
Db	382	AGTGAAAAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGGG	441						
Qy	303	GCATCTCTACTCTCTGGCTGTCTGGGATATGACAGGGGAAAAAGATTGTGTGCAATGAA	362						
Db	442	GCATCTCTACTCTCTGGCTGTCTGGGATATGACAGGGGAAAAAGATTGTGTGCAATGAA	501						
Qy	363	AGTTGTAAAAAGTCCCGACATTATACGGAGACAGCCCTTGGATGAAAATAAAATTGTCAA	422						
Db	502	AGTTGTAAAAAGTCCCGACATTATACGGAGACAGCCCTTGGATGAAAATAAAATTGTCAA	561						
Qy	423	ATGTGTTCCGAAAAGTGCATCCCACTGACCCCAACAAAGACATGCTGCTCCAGCTCATTTGA	482						
Db	562	ATGTGTTCCGAAAAGTGCATCCCACTGACCCCAACAAAGACATGCTGCTCCAGCTCATTTGA	621						
Qy	483	CGATTTCAAAGATTTTCAGGCATGAATGGGATACATGTCTGCAATGCTCTTCAAGTACTTGG	542						
Db	622	CGATTTCAAAGATTTTCAGGCATGAATGGGATACATGTCTGCAATGCTCTTCAAGTACTTGG	681						
Qy	543	CCACATCTCTCAAGTGGATCAATAATCCAACTATCAAGGCTCCAGCTACGTTGCTGTGT	602						
Db	682	CCACATCTCTCAAGTGGATCAATAATCCAACTATCAAGGCTCCAGCTACGTTGCTGTGT	741						
Qy	603	GAAGATGATCATTTCCAGAGTCCCTTCAAGGGTTAGATTACTTACACAGTAAGTCCAAAGT	662						
Db	742	GAAGATGATCATTTCCAGAGTCCCTTCAAGGGTTAGATTACTTACACAGTAAGTCCAAAGT	801						
Qy	663	CATTCAATCTGACATAAAGCCGAAAAATATCTTGAATGTGTGGATGATGATATGTGAG	722						
Db	802	CATTCAATCTGACATAAAGCCGAAAAATATCTTGAATGTGTGGATGATGATATGTGAG	861						
Qy	723	AAAGATGCGAGCTGAGGCCACTGATGTCAGAAAGCAGGTGCTCTCTCTCTCCAGGCTC	782						
Db	862	AAAGATGCGAGCTGAGGCCACTGATGTCAGAAAGCAGGTGCTCTCTCTCTCCAGGCTC	921						
Qy	783	TGCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTTAAAAA	842						
Db	922	TGCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTTAAAAA	981						
Qy	843	AAACTGAAAAAGAAAAACAGAGAGCGAGCTGAGTTATTGAGAGAGCGCTGCGAGGAGT	902						
Db	982	AAACTGAAAAAGAAAAACAGAGAGCGAGCTGAGTTATTGAGAGAGCGCTGCGAGGAGT	1041						
Qy	903	AGAAAGAAATGGAGCGAGAAAGCTGAAAGGAAAAATATAGAGAAAAACATCACCTCAGCTGC	962						
Db	1042	AGAAAGAAATGGAGCGAGAAAGCTGAAAGGAAAAATATAGAGAAAAACATCACCTCAGCTGC	1101						
Qy	963	ACCTTCCAATGACCAAGGATGGCGAATACTGCCCAGAGGTGAAACTTAAAAA	1022						
Db	1102	ACCTTCCAATGACCAAGGATGGCGAATACTGCCCAGAGGTGAAACTTAAAAA	1161						
Qy	1023	AGAGAGAGCGGCTGAGGCGAGAGCTGCAAAAGCAATGCTGAAGTGTAGGACCGAGGAAGA	1082						
Db	1162	AGAGAGAGCGGCTGAGGCGAGAGCTGCAAAAGCAATGCTGAAGTGTAGGACCGAGGAAGA	1221						
Qy	1083	GAAAGAAGATCTGTGAAAGAAAAAATTGAAAAAGATGAAAGATGATGTAGATTCAGGAAC	1142						
Db	1222	GAAAGAAGATCTGTGAAAGAAAAAATTGAAAAAGATGAAAGATGATGTAGATTCAGGAAC	1281						

1143	QY	TGCGAACATAGACCCCTACGTTGGATAGNATCACTCAATAAACCAATGGCCATATTAGAAATGG	1201
1282	DB	TGCGAAACATAGACCCCTACGTTGGATAGNATCACTCAATAAACCAATGGCCATATTAGAAATGG	1341
1203	QY	CCCAATCTCACTGAGCAGCAACCTGGACGATGAAGATGATGATGAAGAAGACATGCCCAAA	1262
1342	DB	CCCAATCTCACTGAGCAGCAACCTGGACGATGAAGATGATGATGAAGAAGACATGCCCAAA	1401
1263	QY	TCCTGAGGAAATATAATCTTGAATGAGCCAAATGAGAAAAGTGAATTAACAATATAGCAGCTC	1322
1402	DB	TCCTGAGGAAATATAATCTTGAATGAGCCAAATGAGAAAAGTGAATTAACAATATAGCAGCTC	1461
1323	QY	CTATGAACAATTAATCAATGGTGAAATGCCAAATGACGACATAAAAATTCGCCAGTCAAGTT	1382
1462	DB	CTATGAACAATTAATCAATGGTGAAATGCCAAATGACGACATAAAAATTCGCCAGTCAAGTT	1521
1383	QY	CCAGAGTTTTTCCACCTCGTTTCTCTGATCCTTAGAACCTGTAGACCTGCGGCTCTCTGT	1442
1522	DB	CCAGAGTTTTTCCACCTCGTTTCTCTGATCCTTAGAACCTGTAGACCTGCGGCTCTCTGT	1581
1443	QY	GCTTTCTGAGGGATCACCACTTACTGAGCAAGAGGAGCAGTCCATCCCATGACAGAAG	1502
1582	DB	GCTTTCTGAGGGATCACCACTTACTGAGCAAGAGGAGCAGTCCATCCCATGACAGAAG	1641
1503	QY	CAGAACGGTTTTACGCTCTCAGTACTCGGGATTTGCCAAAAGCAAAAACCCGGGCACTGTA	1562
1642	DB	CAGAACGGTTTTACGCTCTCAGTACTCGGGATTTGCCAAAAGCAAAAACCCGGGCACTGTA	1701
1563	QY	CTTGTGTGTGAATCCCCTGGATCCGCGGAATGCAGATAAAAATTAGAGTAAAAATTTGCTGA	1622
1702	DB	CTTGTGTGTGAATCCCCTGGATCCGCGGAATGCAGATAAAAATTAGAGTAAAAATTTGCTGA	1761
1623	QY	CCTGGGAAATCTTGTGTGGTGCATAAACACTTCAACGGAAGACAT - CAGACCGGTCAGT	1681
1762	DB	CCTGGGAAATCTTGTGTGGTGCATAAACACTTCAACGGAAGACAT - CAGACCGGTCAGT	1821
1682	QY	ACCGCT - CCATAGAGGTTTTAATAGGACGGGGTACAGCACCCCTCGGACATCTGGAGC	1740
1822	DB	ACCGCTCCCATAGAGGTTTTAATAGGACGGGGTACAGCACCCCTCGGACATCTGGAGC	1881
1741	QY	ACGGCGTGTATGGCAATTTGAGCTGGCAACGGGAGATTAATTTGTTTGAACACACATTTCTGGG	1800
1882	DB	ACGGCGTGTATGGCAATTTGAGCTGGCAACGGGAGATTAATTTGTTTGAACACACATTTCTGGG	1941
1801	QY	GAAGACTAATCCAGAGACGAAGACCAATAGCCCAATCATATAGAGCTGCTAGGCAATTT	1860
1942	DB	GAAGACTAATCCAGAGACGAAGACCAATAGCCCAATCATATAGAGCTGCTAGGCAATTT	2001
1861	QY	CCAAGGCACTTTGCTCTATCTGGAAAATATTTCTCGGAAATTTCTCAATCCGACAGGAGAA	1920
2002	DB	CCAAGGCACTTTGCTCTATCTGGAAAATATTTCTCGGAAATTTCTCAATCCGACAGGAGAA	2061
1921	QY	CTGCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGAAAAAGTAT	1980
2062	DB	CTGCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGAAAAAGTAT	2121
1981	QY	GGCTGGCCCCATGAAGATGCTGCAAGTTTACAGATTTCTTGATCCCGATGTTTGAAGATG	2040
2122	DB	GGCTGGCCCCATGAAGATGCTGCAAGTTTACAGATTTCTTGATCCCGATGTTTGAAGATG	2181
2041	QY	GTTCCAGAAAAACGAGCCCTACGCTGGCGAATGCCTCTGGCATCCTTGTTGAAATCT	2097
2182	DB	GTTCCAGAAAAACGAGCCCTACGCTGGCGAATGCCTCTGGCATCCTTGTTGAAATCT	2238

RESULT 8
AX429244
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

AX429244	3745 bp	DNA	linear	PAT 21-JUN-2002
Sequence	21	from Patent	Epl201765.	

SOURCE	synthetic construct
ORGANISM	synthetic construct
REFERENCE	artificial sequences.
AUTHORS	1
TITLE	Schubart, D., Habenberger, P., Stein-Gerlach, M. and Bevec, D.
JOURNAL	Cellular kinases involved in cytomegalovirus infection and their inhibition
Patent:	EP 1201765-A 21 02-MAY-2002;
FEATURES	Axxima Pharmaceuticals Aktiengesellschaft (DE)
source	Location/Qualifiers
	i. .3745
	/organism="synthetic construct"
	/mol_type="unassigned DNA"
	/db_xref="taxon:32630"
	/note="N/A"
ORIGIN	
	Query Match 95.0%; Score 1991.8; DB 6; Length 3745;
	Best Local Similarity 99.4%; Pred. No. 0;
	Matches 2022; Conservative 0; Mismatches 7; Indels 6; Gaps 2;
Qy	63 TCCGAAAGCCGGAGCCTCAACAGAAAGCTCTTTAGTTCCTCTCTCCACCGCCACC 122
Db	217 TTTAGAAAGCCGGAGCCTCAACAGAAAGCTCTTTAGTTCCTCTCTCCACCGCCACC 276
Qy	123 ACCACACCCAGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGATCCT 182
Db	277 ACCACACCCAGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGATCCT 336
Qy	183 GGGATCAGATGATGAGGAGCAAGAGACCTCTGGGACTACTGCAAGGTGGATATCATCC 242
Db	337 GGGATCAGATGATGAGGAGCAAGAGACCTCTGGGACTACTGCAAGGTGGATATCATCC 396
Qy	243 AGTGAATTTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 302
Db	397 AGTGAATTTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 456
Qy	303 GCACCTTCTACTGTCTGCTGTGGTATGAGGGGAAAGATTTGTTGCAATGAA 362
Db	457 GCACCTTCTACTGTCTGCTGTGGTATGAGGGGAAAGATTTGTTGCAATGAA 516
Qy	363 AGTTGTAAGAGTGGCCAGCATTATACGGAGACAGCTTTGGATGAAATAAAATTTGCTCAA 422
Db	517 AGTTGTAAGAGTGGCCAGCATTATACGGAGACAGCTTTGGATGAAATAAAATTTGCTCAA 576
Qy	423 ATGTGTTTCAGAAAGTGATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA 482
Db	577 ATGTGTTTCAGAAAGTGATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA 636
Qy	483 CGACTTCAAGATTTCAAGGCATGAATGGGATACATGTCGATGGTCTTCAAGTACTTGG 542
Db	637 CGACTTCAAGATTTCAAGGCATGAATGGGATACATGTCGATGGTCTTCAAGTACTTGG 696
Qy	543 CCACCATCTCTCAAGTGATCATCAAAATCCAACTATCAAGGCCTCCCAAGTACTTGGT 602
Db	697 CCACCATCTCTCAAGTGATCATCAAAATCCAACTATCAAGGCCTCCCAAGTACTTGGT 756
Qy	603 GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTACTTACACAGTAAGTCAAGAT 662
Db	757 GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTACTTACACAGTAAGTCAAGAT 816
Qy	663 CATTCATCTGACATAAAGCCGGAATAATCTTTGATGTGTGGATGATGATATGTGAG 722
Db	817 CATTCATCTGACATAAAGCCGGAATAATCTTTGATGTGTGGATGATGATATGTGAG 876
Qy	723 AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCAGGGTC 782
Db	877 AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCAGGGTC 933
Qy	783 TGCAGTGTAGTCCAGTCCACAGCAGAAACCTTAGGAAATAATCTTAAACAAAGAA 842
Db	934 TGCAGTGTAGTCCAGTCCACAGCAGAAACCTTAGGAAATAATCTTAAACAAAGAA 993

Qy	843 AAAAAGCTGAAAAAGAAACAGAGAGGAGGCTGAGTTATTGGAGAACGGCTGCGAGGAT 902
Db	994 AAAAAGCTGAAAAAGAAACAGAGAGGAGGCTGAGTTATTGGAGAACGGCTGCGAGGAT 1053
Qy	903 AGAAGAATTTGGAGCGAGAGCTGAAAGGAAAAATAATAGAGAAAAACATCACCTCAGCTGC 962
Db	1054 AGAAGAATTTGGAGCGAGAGCTGAAAGGAAAAATAATAGAGAAAAACATCACCTCAGCTGC 1113
Qy	963 ACCTTCCATGACACAGGATGGCGAATACTGCCAGAGGTGAAACTTAAAAACAACAGGATT 1022
Db	1114 ACCTTCCATGACACAGGATGGCGAATACTGCCAGAGGTGAAACTTAAAAACAACAGGATT 1173
Qy	1023 AGAGAGCGGCTGAGGAGAGACTGCAAGGACAATGCTGAAGCTGAGGACCCAGGAAGA 1082
Db	1174 AGAGAGCGGCTGAGGAGAGACTGCAAGGACAATGCTGAAGCTGAGGACCCAGGAAGA 1233
Qy	1083 GAAAGAAGTCTGAGAAAGAAAAATTGAAAAAGATGAAGATGATGTAGATCAGAACT 1142
Db	1234 GAAAGAAGTCTGAGAAAGAAAAATTGAAAAAGATGAAGATGATGTAGATCAGAACT 1293
Qy	1143 TCGGAACATAGACCTTACGTGGATAGAATCACCTTAAACCAATGGCCATATTGAGATGG 1202
Db	1294 TCGGAACATAGACCTTACGTGGATAGAATCACCTTAAACCAATGGCCATATTGAGATGG 1353
Qy	1203 CCATTTCTCAGTGGAGCAGCAACTGACGATGAAGATGATGAAGAGACTGCCCAA 1262
Db	1354 CCATTTCTCAGTGGAGCAGCAACTGACGATGAAGATGATGAAGAGACTGCCCAA 1413
Qy	1263 TCCTGAGGAATAAATCTTTGATGAGCCAAATGAGAAAGTGAATACACATATAGCAGCTC 1322
Db	1414 TCCTGAGGAATAAATCTTTGATGAGCCAAATGAGAAAGTGAATACACATATAGCAGCTC 1473
Qy	1323 CTATGAAACAAATTCATGTTGAATTTGCCAAATGGACAGACATAAAATTTCCGAGTCAAGTT 1382
Db	1474 CTATGAAACAAATTCATGTTGAATTTGCCAAATGGACAGACATAAAATTTCCGAGTCAAGTT 1533
Qy	1383 CCAGAGTTTTCACCTGCTGCTCTCTGGAATCTTTAGAACCTGTGCGCTGCGGCTCTGT 1442
Db	1534 CCAGAGTTTTCACCTGCTGCTCTCTGGAATCTTTAGAACCTGTGCGCTGCGGCTCTGT 1593
Qy	1443 GCTTCTGAGGGATCACCACTTTACTGAGCAAGAGGAGCAGTCCATCCATAGACAGAAG 1502
Db	1594 GCTTCTGAGGGATCACCACTTTACTGAGCAAGAGGAGCAGTCCATCCATAGACAGAAG 1553
Qy	1503 CAGAACGGTTTCAGCTTCAGTACTGCGGATTTGCCAAAGCAAAACCCGGGCGAGCTGA 1562
Db	1654 CAGAACGGTTTCAGCTTCAGTACTGCGGATTTGCCAAAGCAAAACCCGGGCGAGCTGA 1713
Qy	1563 CTTGTTGGTGAATCCCTGGATCCCGGAATCGAGATAAAATTTAGAGTAAAAATTCCTGA 1622
Db	1714 CTTGTTGGTGAATCCCTGGATCCCGGAATCGAGATAAAATTTAGAGTAAAAATTCCTGA 1773
Qy	1623 CCTGGAAATGCTTGTGGGTGCATAAAACACTTTACGGAAGACATCCAGACGCGTCAGTA 1682
Db	1774 CTTGGAAATGCTTGTGGGTGCATAAAACACTTTACGGAAGACATCCAGACGCGTCAGTA 1833
Qy	1683 CCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGACCCCTGCGGAACATCTGAGAGCAC 1742
Db	1834 CCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGACCCCTGCGGAACATCTGAGAGCAC 1893
Qy	1743 GCGGTGATGGCATTTGAGCTGGCAACCGGAGATTTTGTGTTGACCCACATTTCTCGGGA 1802
Db	1894 GCGGTGATGGCATTTGAGCTGGCAACCGGAGATTTTGTGTTGACCCACATTTCTCGGGA 1953
Qy	1803 AGACTATTCCAGAGACGAAGACCAATAGCCCAATCATAGAGCTGCTAGGAGTATTTC 1862
Db	1954 AGACTATTCCAGAGACGAAGACCAATAGCCCAATCATAGAGCTGCTAGGAGTATTTC 2013
Qy	1863 AAGGACCTTTGCTCTATCTGGAAAAATTTCTCGGAAATTTCTCAATTCGAGAGGAGAACT 1922
Db	2014 AAGGACCTTTGCTCTATCTGGAAAAATTTCTCGGAAATTTCTCAATTCGAGAGGAGAACT 2073
Qy	1923 GGCACACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGTAAGATATGG 1982

```
Db 2074 GCGACATACCAAGCTGAAGCCCTGAGCCTCTTTGATGCTACTTGTGGAAAAGTATGG 2133
QY 1983 CTGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTGTATCCCGATGTTAGAAATGGT 2042
Db 2134 CTGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTGTATCCCGATGTTAGAAATGGT 2193
QY 2043 TCCAGAAAACGAGCCTCAGCTGGGGAATGCTTCGGCATCCTTGGTTGAATTTCT 2097
Db 2194 TCCAGAAAACGAGCCTCAGCTGGGGAATGCTTCGGCATCCTTGGTTGAATTTCT 2245

RESULT 9
AX666550
LOCUS AX666550 3745 bp DNA linear PAT 26-MAR-2003
DEFINITION Sequence 3 from Patent WO02094796.
ACCESSION AX666550
VERSION AX666550.1 GI:29291032
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Pato,J., Keri,G., Oerfi,L., Waczek,F., Horvath,Z., Banhegyi,P.,
Szabadkai,I., Marosfalvi,J., Hegymegi-Barakonyi,B., sz kelyhidi,Z.,
Greff,Z., Choidas,A., Bacher,G., Daub,H., Obert,S., Kurtenbach,A.,
and Habenberger,P.
TITLE Benz[gl]uinoxaline derivatives as effective compounds against
infectious diseases
JOURNAL Patent: WO 02094796-A 3 28-NOV-2002;
Axixma Pharmaceuticals Aktiengesellschaft (DE)
FEATURES
source
1..3745
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="N/A"

ORIGIN
Query Match 95.0%; Score 1991.8; DB 6; Length 3745;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 2022; Conservative 0; Mismatches 7; Indels 6; Gaps 2;
QY 63 TCGGAAAAGCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCCCTCCACCGCCACC 122
Db 217 TTCAGAAAGCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCCCTCCACCGCCACC 276
QY 123 ACCACCAACCGCCACCTTTGCCAGACCCCAACCCCGGAGCCAGAGGAGAGATCCT 182
Db 277 ACCACCAACCGCCACCTTTGCCAGACCCCAACCCCGGAGCCAGAGGAGAGATCCT 336
QY 183 GGGATCAGATGATGAGGAGCAGAGACCTCGGACTACTGCAAGGTTGATATCATCC 242
Db 337 GGGATCAGATGATGAGGAGCAGAGACCTCGGACTACTGCAAGGTTGATATCATCC 396
QY 243 AGTGAATTTGAGAGCTCTTCAATCGCGGTATCATGTTATTAGAAAGCTTGGATGGGG 302
Db 397 AGTGAATTTGAGAGCTCTTCAATCGCGGTATCATGTTATTAGAAAGCTTGGATGGGG 456
QY 303 GCATTTCTCTACTGCTGTGGCTGTGCTGGGATATGCAAGGGGAAAAGATTTGTCATGAA 362
Db 457 GCATTTCTCTACTGCTGTGGCTGTGCTGGGATATGCAAGGGGAAAAGATTTGTCATGAA 516
QY 363 AGTTCTAAAAGTCCCAGCATTTATCGGAGACAGCCTTGGATGAATAAAATTTGCTCAA 422
Db 517 AGTTGTAAAAGTGGCCAGCATTTATACGGAGACAGCCTTGGATGAATAAAATTTGCTCAA 576
QY 423 ATGTGTTTCGAGAAAGTATCCCAAGTACCACCAACAAAGACATGGTGGTCCAGCTCATTTGA 482
Db 577 ATGTGTTTCGAGAAAGTATCCCAAGTACCACCAACAAAGACATGGTGGTCCAGCTCATTTGA 636
QY 483 CGACTTCAAGATTTTCAGGCGATGAATGGGATACATGCTGCTGCTTTCGAAGTACTTGG 542
```

```
Db 637 CGACTTCAAGATTTTCAGGCGATGAATGGGATACATGCTGTGATGCTCTTTCGAAGTACTTGG 696
QY 543 CCACCATCTCCTCAAGTGGATCATCAATCCAATCATCAAGGCTTCCAGTACCTTGTGTGT 602
Db 697 CCACCATCTCCTCAAGTGGATCATCAATCCAATCATCAAGGCTTCCAGTACCTTGTGTGT 756
QY 603 GAAGAGTATCATTCGACAGCTCTTCAAGGGTTAGATTACTTACACAGTAAAGTCAAGAT 662
Db 757 GAAGAGTATCATTCGACAGCTCTTCAAGGGTTAGATTACTTACACAGTAAAGTCAAGAT 816
QY 663 CATTCATCTGACATTAAGCCGGAATATCTTGTATGCTGTGTGATGATGATGATGATGATGATG 722
Db 817 CATTCATCTGACATTAAGCCGGAATATCTTGTATGCTGTGTGATGATGATGATGATGATGATG 876
QY 723 AAGAAATGGCAGCTGAGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCTCTCTCTCTCT 782
Db 877 AAGAAATGGCAGCTGAGC---CTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCTCTCTCTCT 933
QY 783 TGCAGTCAAGTACGGCTCCACAGCAGAAACCTTATAGGAAAATATCTTAAACAAAAGAA 842
Db 934 TGCAGTCAAGTACGGCTCCACAGCAGAAACCTTATAGGAAAATATCTTAAACAAAAGAA 993
QY 843 AAAACTGAAGAAGAAACAGAGAGCGAGCTGAGTATTGGAGAGCGCCCTGCAGGAGAT 902
Db 994 AAAACTGAAGAAGAAACAGAGAGCGAGCTGAGTATTGGAGAGCGCCCTGCAGGAGAT 1053
QY 903 AGAAGAAATTGAGCGCAGAAAGCTGAAAGGAAAATTAATAGAGAAAACATCACTCAGCTGC 962
Db 1054 AGAAGAAATTGAGCGCAGAGACTGAAAGGAAAATTAATAGAGAAAACATCACTCAGCTGC 1113
QY 963 ACCTTCCAATGACCAGAGTGGCGAATCTGCCAGAGAGTGAACTTAAACAAACAGGATTT 1022
Db 1114 ACCTTCCAATGACCAGAGTGGCGAATCTGCCAGAGAGTGAACTTAAACAAACAGGATTT 1173
QY 1023 AGAGAGCGCGCTGAGCGAGAGACTGCAAGAGCAATGGTGAAGCTGAGGACCCAGGAGAA 1082
Db 1174 AGAGAGCGCGCTGAGCGAGAGACTGCAAGAGCAATGGTGAAGCTGAGGACCCAGGAGAA 1233
QY 1083 GAAAGAAAGTCTGTAGAAAGAAAACATTTGAAAAAGATGAAGATGATGATGATGATGATGATGATG 1142
Db 1234 GAAAGAAAGTCTGTAGAAAGAAAACATTTGAAAAAGATGAAGATGATGATGATGATGATGATG 1293
QY 1143 TCGGAAACATAGACCTTACGTGATAGATCACTTAACCAATGATGATGATGATGATGATGATGATG 1202
Db 1294 TCGGAAACATAGACCTTACGTGATAGATCACTTAACCAATGATGATGATGATGATGATGATGATG 1353
QY 1203 CCCATTCTCACTGGAGCAGCAACTGGACGATGAAGATGATGATGATGATGATGATGATGATGATG 1262
Db 1354 CCCATTCTCACTGGAGCAGCAACTGGACGATGAAGATGATGATGATGATGATGATGATGATGATG 1413
QY 1263 TCTGAGGAATATAATCTTGTATGAGCCAAATGCGAGAAAGTATTAACATATAGCAGCTC 1322
Db 1414 TCTGAGGAATATAATCTTGTATGAGCCAAATGCGAGAAAGTATTAACATATAGCAGCTC 1473
QY 1323 CTATGAACAAATTCATGTTGAAATTCGCCAAATGCGACATATAAATTCGCGAGTCAAGTT 1382
Db 1474 CTATGAACAAATTCATGTTGAAATTCGCCAAATGCGACATATAAATTCGCGAGTCAAGTT 1533
QY 1383 CCCAGAGTTTTCACCTCGTTGTTCTCTGAGATCTTTAGAACCTGTGGCTGCGGCTCTGT 1442
Db 1534 CCCAGAGTTTTCACCTCGTTGTTCTCTGAGATCTTTAGAACCTGTGGCTGCGGCTCTGT 1593
QY 1443 GCTTTCTGAGGATCAACACTTACTGAGCAAGAGAGAGCAGTCCATCCCATGACAGAG 1502
Db 1594 GCTTTCTGAGGATCAACACTTACTGAGCAAGAGAGAGCAGTCCATCCCATGACAGAG 1653
QY 1503 CAGAACGGTTTTCAGCTCAGTACCTGCGGATTTGCCAAAGCAAAACCCGCGCAGCTGA 1562
Db 1654 CAGAACGGTTTTCAGCTCAGTACCTGCGGATTTGCCAAAGCAAAACCCGCGCAGCTGA 1713
QY 1563 CTTGTTGTGTAATCCCTCGATCCGCGAATCCAGATATAAATTAGAGTAAATAATTTGCTGA 1622
```

Db	1714	CTTTGGTGGTAATCCCCTGGATCCGCGGAATCGAGATAAAATTAGAGTAAAAATTCGTGA	1773
Qy	1623	CCTCGGAAATGCTTTGTTGGGTGCATAAACACTTCCGGAAGACATCCAGACGCGTCAAGTA	1682
Db	1774	CCTCGGAAATGCTTTGTTGGGTGCATAAAACACTTCACGGAAGACATCCAGACGCGTCAAGTA	1833
Qy	1683	CCGCTCCATAGAGGTGTTTAATAGAGCGGGGTACAGCACCCCTCGCGACATCTCGAGCAC	1742
Db	1834	CCGCTCCATAGAGGTGTTTAATAGAGCGGGGTACAGCACCCCTCGCGACATCTCGAGCAC	1893
Qy	1743	GGCGTGATGGCATTTGAGCTGGCAACGGGAGATTATTGTTGTGAACCAATTCCTGGGGA	1802
Db	1894	GGCGTGATGGCATTTGAGCTGGCAACGGGAGATTATTGTTGTGAACCAATTCCTGGGGA	1953
Qy	1803	AGACTATTCAGAGACGGAACACCATAGCCCATCATAGAGCTGTAGCGAGTATTC	1862
Db	1954	AGACTATTCAGAGACGGAACACCATAGCCCATCATAGAGCTGTAGCGAGTATTC	2013
Qy	1863	AAGGCATTTGCTCTATCTCGGAAATATTCTCGGGAATTTCTTCAATCGCAGAGAGAACT	1922
Db	2014	AAGGCATTTGCTCTATCTCGGAAATATTCTCGGGAATTTCTTCAATCGCAGAGAGAACT	2073
Qy	1923	CGGACATCATCAAGCTGAAGCCCTCGAGCCTCTTTGATGTACTTTGTGAAAAGTATGG	1982
Db	2074	CGGACATCATCAAGCTGAAGCCCTCGAGCCTCTTTGATGTACTTTGTGAAAAGTATGG	2133
Qy	1983	CTGSCCCCATGAAGATGCTGCACAGTTTATACAGATTTCTCGATCCGATGTTTAGAAAATGGT	2042
Db	2134	CTGSCCCCATGAAGATGCTGCACAGTTTATACAGATTTCTCGATCCGATGTTTAGAAAATGGT	2193
Qy	2043	TCGAAAAAAGAGCCCTCAGCTGCGGAATGCTTTCGGCATCCCTTGGTTGAATCT	2097
Db	2194	TCGAAAAAAGAGCCCTCAGCTGCGGAATGCTTTCGGCATCCCTTGGTTGAATCT	2245

```

RESULT 10
HSU88666
LOCUS
DEFINITION Homo sapiens serine kinase SRPK2 mRNA, complete cds.
ACCESSION U98666
VERSION U98666.1 GI:1857943
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3745)
Wang,H.Y., Lin,W., Dyck,J.A., Yeakley,J.M., Songyang,Z.,
SPK2,L.C. and Fu,X.D.
SRPK2: a differentially expressed SR protein-specific kinase
involved in mediating the interaction and localization of pre-mRNA
splicing factors in mammalian cells
J. Cell Biol. 140 (4), 737-750 (1998)
98139536
MEDLINE
PUBMED 9472028
2 (bases 1 to 3745)
Wang,H.Y., Wen,L. and Fu,X.D.
Direct Submission
Submitted (06-FEB-1997) Cell. Mol. Medicine, University of
California at San Diego, 9500 Gilman Drive, La Jolla, CA
92093-0651, USA
FEATURES
source
Location/Qualifiers
1..3745
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
188..2248
/note="similar to human serine kinase SRPK1, encoded by
GenBank Accession Number U09564; specific for
serine/arginine-rich splicing factors"
/codon_start=1
/product="serine kinase SRPK2"
/protein_id="AAC05299.1"
CDS

```

994	Db	GAAAAAGAAA CAGAAGAGGCGAGGCTGAGTTATTGGAGAAGCGCCTGCAGGAGAT	1053
903	Qy	AGAAGATTGGACGGAGAACTGAAAGGAATAATAGTAGAAGAAACATCACCTCAGCTGC	962
1054	Db	AGAAGAAATTGGACGGAGAACTGAAAGGAATAATAGAGAAGAAACATCACCTCAGCTGC	1113
963	Qy	ACCTTCCAAATGACCAGGATGCGCAATACTGCCACAGAGTCAAACTAAAAACAACAGGATT	1022
1114	Db	ACCTTCCAAATGAACAGGATGCGCAATACTGCCACAGAGTGAACTAAAAACAACAGGATT	1173
1023	Qy	AGAGGAGCGCGCTGAGGCACAGACTGCAAAAGGAACAATGGTGAAGCTGAGGAACAGGAAGA	1082
1174	Db	AGAGGAGCGCGCTGAGGCACAGACTGCAAAAGGAACAATGGTGAAGCTGAGGAACAGGAAGA	1233
1083	Qy	GAAGAAGATGCTGAGAGAAGAAAACATTGAAAAGATGAAGATGATGTAGATCAGGAAC	1142
1234	Db	GAAGAAGATGCTGAGAGAAGAAAACATTGAAAAGATGAAGATGATGTAGATCAGGAAC	1293
1143	Qy	TGCGAACATAGACCTACGTGGATAGAAATCACCTAAAAACCAATGGCCATATTGAGAATGG	1202
1294	Db	TGCGAACATAGACCTACGTGGATAGAAATCACCTAAAAACCAATGGCCATATTGAGAATGG	1353
1203	Qy	CCCAATCTCACTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAAGATGCCCAA	1262
1354	Db	CCCAATCTCACTGGAGCAGCAACTGGACGATGAAGATGATGATGAAGAAGATGCCCAA	1413
1263	Qy	TCCTGAGGAATATTAATCTTGATGAGCCAAATGCGAAAGTGATTACATATATAGCAGCTC	1322
1414	Db	TCCTGAGGAATATTAATCTTGATGAGCCAAATGCGAAAGTGATTACATATATAGCAGCTC	1473
1323	Qy	CTATGAACAATTCAATGSGTGAATTGCCAAATGGACGACATAAAAATTCGCCAGTCA	1382
1474	Db	CTATGAACAATTCAATGSGTGAATTGCCAAATGGACGACATAAAAATTCGCCAGTCA	1533
1383	Qy	CCACAGATTTTCCACCTCGTTGTTCTCTGGAATCCTTAGAAACCTGTGSCCTGCGGCTCTGT	1442
1534	Db	CCACAGATTTTCCACCTCGTTGTTCTCTGGAATCCTTAGAAACCTGTGSCCTGCGGCTCTGT	1593
1443	Qy	GCTTTCTGAGGGATCACACTTACTGAGCAAGAGGAGAGCAGTCCATCCATGACAGAAG	1502
1594	Db	GCTTTCTGAGGGATCACACTTACTGAGCAAGAGGAGAGCAGTCCATCCATGACAGAAG	1653
1503	Qy	CAGAACGGTTTCAGCCTCCAGTACTGGGGATTTCGCCAAAAGCAAAAACCCGGGCGAGCTGA	1562
1654	Db	CAGAACGGTTTCAGCCTCCAGTACTGGGGATTTCGCCAAAAGCAAAAACCCGGGCGAGCTGA	1713
1563	Qy	CTTGTGTGTAATCCCTTGGATCCGCGGAATGCAAGATAAAATTAGAGTAAAAATTTGCTGA	1622
1714	Db	CTTGTGTGTAATCCCTTGGATCCGCGGAATGCAAGATAAAATTAGAGTAAAAATTTGCTGA	1773
1623	Qy	CCTGGGAAATGCTTTGTTGGTGCATAAACCTTCAACGGAAGACATCCAGACGGTCAAGTA	1682
1774	Db	CCTGGGAAATGCTTTGTTGGTGCATAAACCTTCAACGGAAGACATCCAGACGGTCAAGTA	1833
1683	Qy	CCGCTCCATAGAGTTTTAATAGAGCGGGGTACAGCACCCCTCGCGACATCTGGAGCAC	1742
1834	Db	CCGCTCCATAGAGTTTTAATAGAGCGGGGTACAGCACCCCTCGCGACATCTGGAGCAC	1893
1743	Qy	GGCGTGTATGGCAATTGAGCTGGCAACGGGAGATTATTGTTTGAACCCACATCTCGGGGA	1802
1894	Db	GGCGTGTATGGCAATTGAGCTGGCAACGGGAGATTATTGTTTGAACCCACATCTCGGGGA	1953
1803	Qy	AGACTATTCACAGACGGAAGACACATAGCCCAACATCATAGAGCTGCTAGCAGTATTC	1862
1954	Db	AGACTATTCACAGACGGAAGACACATAGCCCAACATCATAGAGCTGCTAGCAGTATTC	2013
1863	Qy	AAGGCACTTTCCTTATCTGGAATAATTCTCGGGAATTTCTCAATCGCAGAGGAGAACT	1922
2014	Db	AAGGCACTTTCCTTATCTGGAATAATTCTCGGGAATTTCTCAATCGCAGAGGAGAACT	2073
1923	Qy	CGCACATCATCAAGCTGGAAGCCCTGGAGCCTCTTTGATGTAATCTTGGAAGAAAGTATGG	1982

D	b	2074	GGGACATCACCAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTGTGGAAGAATGTCG	2133
Q	y	1983	CTGGCCCCCATGAAGATGCTGCACAGTGTTCACAGATTTTCAGATTTCCCGATCCCGATGTTAGAAATGGT	2042
D	b	2134	CTGGCCCCCATGAAGATGCTGCACAGTGTTCAGATTTCCCGATCCCGATGTTAGAAATGGT	2193
Q	y	2043	TCCAGAAAAACGAGCCTCAGCTGGCGAATGCCCTTCGGCATCCTTGTTGTAATTCT	2097
D	b	2194	TCCAGAAAAACGAGCCTCAGCTGGCGAATG---TCGGCATCCTTGTTGTAATTCT	2245
 RESULT 11 AY354201 3028 bp mRNA linear PRI 02-SEP-2003 LOCUS Homo sapiens SFRS protein kinase 2 isoform c (SRPK2) mRNA, complete cds, alternatively spliced. ACCESSION AY354201 VERSION AY354201.1 GI:34329349 KEYWORDS "Homo sapiens (human)" SOURCE Homo sapiens (human) ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE AUTHORS Sha,J.H., Zhou,Z.M. and Li,J.M. TITLE Direct Submission JOURNAL Submitted (29-JUL-2003) Key Lab of Reproductive Medicine, Nanjing Medical University, 140 Han Zhong Road, Nanjing, Jiangsu 210029, China FEATURES Location/Qualifiers 1..3028 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /chromosome="7" /tissue_type="testis" 1..3028 /gene="SRPK2" 223..1863 /note="transcript variant 3; alternatively spliced" /codon_start=1 /product="SFRS protein kinase 2 isoform c" /protein_id="AA063886.1" /db_xref="GI:34329350" /translation="MSVNSEKSSSERPEPQQAKLVPVPPPPPPPPLPDPTPEE PEEILGSDDEEDPADYCKGYHPVGIDLFNGRYHVIRKLGNWHFSTVWLMDMG GKRFVAVGHVKIAHYETALDEIKLLKCVRKESDPSDFNKMVMVQLIIDPFKISGMGIG HVCVFVGLVHLKIITIKSNYQGLPVRCYSPIRVOGLDYLSHKCKIHTDDIKPE NILDMQVLRBAABATBOKAGAPPSPGSASTAPQOKIPGKTISKNNKKLKKKQ KROALLEKLRIIEELERAEARKITEENTSAAPNDQGEYCPEVKLTGTGLEAAA EAETAKNGAEOQEKEADEKENIEKDEDVDQELANIPTWIESPTKNTGILENGFP SLEQLDDEDDDCPNREINLDENAESDTYSSSYEQFNGEFLPNGRHKKIPESQF PEFTSLFDGSLFPVACGVSLSGSLTEQSEESSPHDRSRRTVSASSTGDLPKAKTRA ADLVNPLDPNRADKRVKIADIAGLNACVWQKPQLTCW"				
 ORIGIN				
Query Match 86.5%; Score 1814.6; DB 9; Length 3028;				
Best Local Similarity 91.1%; Pred. No. 0;				
Matches 2031; Conservative 0; Mismatches 4; Indels 194; Gaps 2;				
Q	y	63	TCCGAAAAAGCGCGAGCCTCAACAGAAAGCTCTCTTTAGTTCTCTCTCTCACCGCCACC	122
D	b	252	TTCAGAAAGCGCGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCACCGCCACC	311
Q	y	123	ACCACCAACCGCCACCTTTGCCAGACCCCACACCCCGGAGCCAGAGGAGGATCCT	182
D	b	312	ACCACCAACCGCCACCTTTGCCAGACCCCACACCCCGGAGCCAGAGGAGGATCCT	371
Q	y	193	GGGATCATGATGAGGAGCAAGAGGACCTCGCGACTACTGCAAAGGTGGATATCATCC	242
D	b	372	GGGATCATGATGAGGAGCAAGAGGACCTCGCGACTACTGCAAAGGTGGATATCATCC	431

Qy	243	AGTGAATAATTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG	302
Db	432	AGTGAATAATTGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG	491
Qy	303	GCATCTCTACTGCTGTGGCTGTGCTGGGATATGACAGGGGAAAGATTTGTTGCAATGAA	362
Db	492	GCATCTCTACTGCTGTGGCTGTGCTGGGATATGACAGGGGAAAGATTTGTTGCAATGAA	551
Qy	363	AGTTGTAATAAGTGGCCAGCATTTATACGGAGACAGCCTTGGATGAAATAAAATTCGCTCAA	422
Db	552	AGTTGTAATAAGTGGCCAGCATTTATACGGAGACAGCCTTGGATGAAATAAAATTCGCTCAA	611
Qy	423	ATGTGTTTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGTGTGTGTCAGCTCATTTGA	482
Db	612	ATGTGTTTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGTGTGTGTCAGCTCATTTGA	671
Qy	483	CGACTTCAAGATTTTCAGCAGTGAATGGGATACATGCTGCAATGCTCTTCGAAGTACTTGG	542
Db	672	CGACTTCAAGATTTTCAGCAGTGAATGGGATACATGCTGCAATGCTCTTCGAAGTACTTGG	731
Qy	543	CCACCATCTCTCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCCTCCAGTAGTCTGTGT	602
Db	732	CCACCATCTCTCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCCTCCAGTAGTCTGTGT	791
Qy	603	GAAGAGTATCATTTTCGACAGGTCTCTTCAAGGGTTAGATTACTTTACACAGTGAAGTCAAGAT	662
Db	792	GAAGAGTATCATTTTCGACAGGTCTCTTCAAGGGTTAGATTACTTTACACAGTGAAGTCAAGAT	851
Qy	663	CATTCACTGACATAAAGCCGGAATAATCTTCGATGTGTGATGATGCTATGCTATGTGAG	722
Db	852	CATTCACTGACATAAAGCCGGAATAATCTTCGATGTGTGATGATGCTATGCTATGTGAG	911
Qy	723	AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCTCTTCAGGGTC	782
Db	912	AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCTCTTCAGGGTC	971
Qy	783	TGCAGTGAGTACGGCTCCACAGCAGAAAACCTATAGGAAAAATATCTTAAAAACAAAAAGAA	842
Db	972	TGCAGTGAGTACGGCTCCACAGCAGAAAACCTATAGGAAAAATATCTTAAAAACAAAAAGAA	1031
Qy	843	AAATCTGAAAAAGAAAACAGAGAGCGCAGGTGAGTTATTGGAGAGCGCCTGACAGGAGAT	902
Db	1032	AAATCTGAAAAAGAAAACAGAGAGCGCAGGTGAGTTATTGGAGAGCGCCTGACAGGAGAT	1091
Qy	903	AGAAAGATTTGGAGCGCAGAGCTGAAAGGAAAAATATAGAGAAAAACATCACCTCAGCTGC	962
Db	1092	AGAAAGATTTGGAGCGCAGAGCTGAAAGGAAAAATATAGAGAAAAACATCACCTCAGCTGC	1151
Qy	963	ACCTTCCAATGACAGGATGGCGAATATGCCCCAGAGGTGAAACTTAAAAACAAACAGGAT	1022
Db	1152	ACCTTCCAATGACAGGATGGCGAATATGCCCCAGAGGTGAAACTTAAAAACAAACAGGAT	1211
Qy	1023	AGAGGAGCGGCTGAGCGAGAGACTGCMAAGGACAATGGTGAAGCTGAGGACCGAGGAAGA	1082
Db	1212	AGAGGAGCGGCTGAGCGAGAGACTGCMAAGGACAATGGTGAAGCTGAGGACCGAGGAAGA	1271
Qy	1083	GAAAGAGAATGCTCAGAGAAAGAAAACATTTGAAAAAGATGCAAGATGATGTAGATCAGGAAC	1142
Db	1272	GAAAGAGAATGCTCAGAGAAAGAAAACATTTGAAAAAGATGCAAGATGATGTAGATCAGGAAC	1331
Qy	1143	TGCGAAACATAGACCTCAGTGGATAGAAATCACTTAAAAACCAATGGCCCATTTTGAGATGG	1202
Db	1332	TGCGAAACATAGACCTCAGTGGATAGAAATCACTTAAAAACCAATGGCCCATTTTGAGATGG	1391
Qy	1203	CCCATTTCTACTGAGCAGCAGCACTGGACGATGAAGATGATGATGAAGAGACTGCCCAA	1262
Db	1392	CCCATTTCTACTGAGCAGCAGCACTGGACGATGAAGATGATGATGAAGAGACTGCCCAA	1451
Qy	1263	TCCTGAGGAATATAATCTTGATGAGCCAAATGACAGAAAGTGATTTACACATATACAGCTC	1322
Db	1452	TCCTGAGGAATATAATCTTGATGAGCCAAATGACAGAAAGTGATTTACACATATACAGCTC	1511
Qy	1323	CTATGAAACAAATTCATGGTGAAATTCGCCAAATGGACGACATAAAAATTCGCCAGTACAGTT	1382

D	b		1512	CTATGAAACAATTCAATGGTGAATTCGCAMAATGGACGACATATAAATTTCCAGAGTCACAGTT	1571
Q	y		1383	CCCAGAGTTTTCCACCTCGTTGTCTCTGGACTCTTAGAACCTGTGTGCCCTCGCGCTCTGT	1442
D	b		1572	CCCAGAGTTTTCCACCTCGTTGTCTCTGGACTCTTAGAACCTGTGTGCCCTCGCGCTCTGT	1631
Q	y		1443	GCTTTCTGAGGGATFCAACACTTA CTAGCAGAAGAGGAGAGCAGTGCCATGCCATGACAGAAG	1502
D	b		1632	GCTTTCTGAGGGATFCAACACTTA CTAGCAGAAGAGGAGAGCAGTGCCATGCCATGACAGAAG	1691
Q	y		1503	CAGAACGGTTTTACGGCTTCAGTACTGCGGGATTTTGCCAAAA-	1542
D	b		1692	CAGAACGGTTTTACGGCTTCAGTACTGCGGGATTTTGCCAAAAACCGCGGACGTGA	1751
Q	y		1543	-	1542
D	b		1752	CTTGTGTGTGAATCCCCCTGGATCCGCGGAATGCAGATAAAATTAGAGTAAAAATTTGCTGA	1811
Q	y		1543	-	1581
D	b		1812	CCTGGGAAATGCTTGTGGGTGCAAAAACCCGGGACGTGACTGTGTGGTGAATCCCCCTG	1871
Q	y		1582	GATCCGCGGAATGCAGATAAAAATTTAGAGTAAAAATTTGCTGACCTGGGAAATGCTGTGTGG	1641
D	b		1872	GATCCGCGGAATGCAGATAAAAATTTAGAGTAAAAATTTGCTGACCTGGGAAATGCTGTGTGG	1931
Q	y		1642	GTGCATAAACACTTTCACGGAAGACATCCAGACGGTCAAGTACCGTCCATAGAGGTTTTA	1701
D	b		1932	GTGCATAAACACTTTCACGGAAGACATCCAGACGGTCAAGTACCGTCCATAGAGGTTTTA	1991
Q	y		1702	ATAGCAGCGGGTACAGCACCCCTGCGGACATCTGGAGCAGCGGCTGTATGGCATTTTGAG	1761
D	b		1992	ATAGCAGCGGGTACAGCACCCCTGCGGACATCTGGAGCAGCGGCTGTATGGCATTTTGAG	2051
Q	y		1762	CTGGCAAACGGGAGATTAATTTGTTGAAACCACTTCGGGGAAGACTATTCAGAGACGAA	1821
D	b		2052	CTGGCAAACGGGAGATTAATTTGTTGAAACCACTTCGGGGAAGACTATTCAGAGACGAA	2111
Q	y		1822	GACCACATAGCCACATCATAGAGCTCTAGGAGTATTCAGAGGCACTTTGCTCTATCT	1881
D	b		2112	GACCACATAGCCACATCATAGAGCTCTAGGAGTATTCAGAGGCACTTTGCTCTATCT	2171
Q	y		1882	GGAAAAATTTCTCGGGAAATTTCTCAATTCGAG-	1913
D	b		2172	GGAAAAATTTCTCGGGAAATTTCTCAATTCGAGAGATCACATAGCATTTGATGAACCTG	2231
Q	y		1914	-	1913
D	b		2232	CTGGGGAAGTCCCTCGAAAATAGCTATGTTGGGGAATFACTCCAAGAGTTTTTCACC	2291
Q	y		1914	-----AGGAGAACTCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGACTTT	1968
D	b		2292	AGAAAAGGAGAACTCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGACTTT	2351
Q	y		1969	GTGAAAAGTATGCTGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTTGATCCCG	2028
D	b		2352	GTGAAAAGTATGCTGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTTGATCCCG	2411
Q	y		2029	ATGTTAGAAATGTTTCCAGAAAAACGACCTCAGCTGGCGAATGCCCTTCGGCATCTCTGG	2088
D	b		2412	ATGTTAGAAATGTTTCCAGAAAAACGACCTCAGCTGGCGAATGCCCTTCGGCATCTCTGG	2471
Q	y		2089	TTGAAATCTT 2097	
D	b		2472	TTGAAATCTT 2480	

RESULT 12
AB006036
LOCUS
DEFINITION
ACCESSION

AB006036 2300 bp mRNA linear ROD 25-DEC-1997
Mus musculus mRNA for SRPK2, complete cds.
AB006036

VERSION	AB060036.1	GI:2723281
KEYWORDS	SRPK2.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	1 (sites)	
TITLE	Kuroyanagi,N., Onogi,H., Wakabayashi,T. and Hagiwara,M.	
JOURNAL	Novel SR-protein-specific kinase,SRPK2,diseassembles nuclear speckles	
REFERENCE	Biochem. Biophys. Res. Commun. (1998) In press	
AUTHORS	2 (bases 1 to 2300)	
TITLE	Hagiwara,M.	
JOURNAL	Direct Submission	
COMMENT	Submitted (25-JUL-1997) Masatoshi Hagiwara, Tokyo Medical and Dental University Research Institute, Department of Endocrinology; Yushima, Bunkyo-ku, Tokyo 113, Japan (E-mail:m.hagiwara.end@mri.tmd.ac.jp. Tel:03-5803-5836, Fax:03-5803-5836)	
FEATURES	Sequence updated (16-DEC-1997).	
Source	Location/Qualifiers	
	1..2300	
	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/db_xref="taxon:10090"	
CDS	1..2046	
	/codon_start=1	
	/product="SRPK2"	
	/protein_id="BAA24055.1"	
	/db_xref="GI:2723282"	
	/translation="MSVNSESSSSERPEPOQKAPLVPPPPPPPPPLDPAPPEPE RTFLGSDDEQDPADYCKGGYHPVKIGDLNFRYIRKLGWGHSTVWLCDMGGK EFVAMKVSQAQHYETALDEIKLLKCVRESDPDNKMVQLIDDPKISGMDGHHV CMWPEVLGHLLKWLIIKSNQGLPVRCVKSIIROVLQGLDYLKSKKIIHTDIKPNII LMCVDLVYRMAAATSWQAGAPPSSGSAVTAPOQPIGKISKKKKLLKKQKR QALLKRRQLEIELEERAEERKILENITSAAEQDGEVQPEVTLKAADLEDTTEE ETAKNGVEDDEEDAEKNAEKDDVDQEPANLDPTVWESPKNAGHENGPSL EQLEEDDEDDCANPEYNLDNAESDVTYSSEYQFNGLPQHQKTSFPTPL FSGPLPVAKVSISEGPLTQBESSPSHDSRTVSASSTGDLPKTKRAADLLNVP LDPRNADKRVKIDALGNACWVHKHFTDIOTROYRSIEVLIGAGYSTPADINSTACM AFELATGDLVLEPHSGEDYSRDEHIAHIELLGSIPIHFALSGKYSEPFNRRGELR HITKLKPSLIFDLVLEKYGWPHEDAAQFTDFLIPMLEVMPKPRASAGECLRHPLWLS" 2300	
	/note="18 A nucleotides"	
polyA_eite		
ORIGIN		
Query Match	79.3%;	Score 1662.6; DB 10; Length 2300;
Best Local Similarity	89.4%;	Pred. No. 0;
Matches 1820;	Conservative	0; Mismatches 194; Indels 21; Gaps 2;
QY	63	TCCGAAAAGCCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACGCCACC 122
DB	30	TTGAGAAAGCCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACGCCACC 89
QY	123	ACCCACCCACCCGACCTTTGCGACACCCACACCCCGGAGCCAGGAGGAGATCCT 182
DB	90	ACCGCCACCCAC-----TGCAGACCCCGCGCCCCCAGAGCCAGGAGGAGATTCT 143
QY	183	GGGATCAGATGATGAGGAGCAAGGAGACCTCGGGACTACTGCAAGAGTGATATCATCC 242
DB	144	GGGGTCAGATGATGAGGAGCAGGAGGAGCCCGCAGATTACTGCAAGAGTGCGTATCATCC 203
QY	243	AGTGAATAATTGAGACCTCTCAATGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 302
DB	204	AGTGAATAATTGAGATCTCTTCAATGGTCGATCATGTCATTAGAAAGCTAGGATGGG 263
QY	303	GCACTTCTCTACTGTCGGCTGTGCTCGGGATATGCGGGGAAAGAAGATTGTTGCAATGAA 362
DB	264	GCACTTTTCTACTGTATGGCTGTGCTCGGGATATGCGGGGAAAGAAGATTGTTGCAATGAA 323
QY	363	AGTTGTAAAGAGTCCCGAGCATTTATACGGAGACAGCCTTGGATGAATAAATTTGCTCAA 422
DB	324	AGTTGTAAAGAGTCCCGAGCATTTATACGGAGACAGCCTTGGATGAATAAATTTAACTCTCA 383

Best Local Similarity 89.5%; Pred. No. 0;			
Matches 1825; Conservative 0; Mismatches 189; Indels 24; Gaps 3;			
Qy	63	TCCGAAAGCGGAGCCTCAACGAAAGCTCTTTAGTTCCTCCTCCTCCACCGCCACC	122
Db	134	TTCAAGAAAGCGGAGCCTCAACGAAAGCTCTTTAGTTCCTCCTCCTCCACCGCCACC	193
Qy	123	ACCACACACCGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGAGATCCT	182
Db	194	ACCGCACACAC-----TGCCAGACCCCGCACCCCGAGCCAGAGGAGAGATCT	247
Qy	183	GGATCAGATGATGAGGAGCAAGAGACCTTGCGGACTATGCGAAGGTGATATCATCC	242
Db	248	GGGCTCAGATGATGAGGAGCAAGAGACCCCGCAGATTTACTGCAAGGTGCTATCATCC	307
Qy	243	AGTGAAGATGAGACCTCTTCAATGGCGGATATCATGTTTGAAGAGCTTGGATGGG	302
Db	308	AGTGAAGATGAGAGATCTCTTCAATGGCTCGATATCATGTCATTAAGAAAGCTAGAGTGGG	367
Qy	303	GCATCTCTACTGTCTGCTGTGCTGGGATATGAGGGGAAAGATTTGTGCAATGAA	362
Db	368	GCATCTTCTACTGTATGGCTGTGCTGGATATGCAAGGAAAGATTTGTGCAATGAA	427
Qy	363	AGTTGTAAAGTGCCTCAGCATATATACGAGACAGCCTTGGATGAAATTAATTTGCTCAA	422
Db	428	AGTTGTAAAGTGCCTCAGCATATATACGAGACAGCCTTGGATGAAATTAATTTGCTCAA	487
Qy	423	ATGTGTCGAGAAAGTGATCCAGTGACCCCAACCAAGACATGGTGGTCCAGCTCATTTGA	482
Db	488	ATGCGCTCGAGAAAGTGATCCAGTGACCCCAACCAAGACATGGTGGTCCAGCTCATTTGA	547
Qy	483	CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCAATGCTCTCGAAGTACTTGG	542
Db	548	TGACTTCAAGATCTCAGGCATGAATGGGATACATGTCTGCAATGCTCTCGAAGTACTTGG	607
Qy	543	CCACCATCTCTCAAGTGATCATCAATCCAACTATACAGGCCCTCCAGTACTTGTGT	602
Db	608	TCAACCATCTCTCAATGATCATCAATCCAACTATACAGGCCCTCCAGTACTTGTGT	667
Qy	603	GAGAGATCATTTCCAGACAGTCTCTCAAGGTTTGAATTTACACAGTAAGTGCAAGAT	662
Db	668	GAGAGATCATTTCCAGACAGTCTCTCAAGGTTTGAATTTACACAGTAAGTGCAAGAT	727
Qy	663	CATTCATCTGACATAAAGCCGAAATATCTTGATGTGTGTGGATGATGCAATATGTGAG	722
Db	728	AATTCACCGACATAAAGCCGAAATATCTTGATGTGTGTGGATGATGCAATATGTGAG	787
Qy	723	AAGATGGCAGCTGAGGCCCATGATGGCAGAAAGCAGAGTGTCTCTCTCTCAGGCTC	782
Db	788	AAGATGGCAGCCGAGCCAGCAGAGTGGCAGAAAGCAGAGTGTCTCTCTCTCAGGCTC	847
Qy	783	TGCAGTGATGAGCTCCACAGCAGAAACCTATAGGAAATATCTTAAAAACAAAAAGAA	842
Db	848	TGCAGTGATGAGCTCCACAGCAAAACCTATAGGAAATATCTTAAAAACAAAAAGAA	907
Qy	843	AAACTGAAAAAGAAACAGAGAGGAGCAGCTGAGTTTATTTGGAGAGCGCTCGAGAGAT	902
Db	908	AAAGCTGAAAAAGAAACAGAGAGAGCAGCTGAGTTTATTTGGAGAGCGCTCGAGAGAT	967
Qy	903	AGAGAAATTTGGACGAGAGCTGAAAGGAAATATATAGAGAAACATCACCTCAGCTGC	962
Db	968	TGAGGAATTTGGACGAGAGCCGAAAGGAAATCTCTAGAGGAGAACATCACCTCTCGAGA	1027
Qy	963	ACCTTCC---AATGACAGAGTGGCAATACTGCCAGAGGTGAAACTTAAAAACACACAGG	1019
Db	1028	AGCTTCGGGGACGACGAGATGGAGATACCAAGCCGAGGTGACACTGAAGACCCGCA	1087
Qy	1020	ATTAGAGGAGCGGCTGAGCAGAGACTGCAAAAGGACAAATGGTGAAGCTGAGACACAGGA	1079
Db	1088	CTTAGAGGACACAACTGAGGAGAGACAGCAAAAGGATAATGGTGAAGTTGAAGACACAGGA	1147
Qy	1080	AGAGAAAGATGCTTGAGAGAGAAACATTTGAAAGATGAGATGATCTAGATCAGGA	1139

Db	1148	AGAGAAAGATGTCAGAGAGGAGAAACGCGGAGAGGATGAAGATGATGTGAACAGGA	1207
Qy	1140	ACTTGGAAACATAGACCTTACGTGGATAGAAATCACCTAAACCAATGGCCATATTAGAA	1199
Db	1208	ACTTGGAAACCTTAGACCTTACGTGGTGGAGTCCCCGAAAGCCAAATGGCCATATTGAAA	1267
Qy	1200	TGGCCCATTTCTACTTGGAGCAGCAACTGGACGATGAGATGATGATGAAGAGACTGCC	1259
Db	1268	TGGCCCATTTCTACTTGGAGCAGCAGCTGGAGGATGAAGAGGAGATGAAGATGACTGTGC	1327
Qy	1260	AAATCTGAGGAATATATCTTGTATGAGCCAAATGCAAGAAAGTGAATTACATATAGCAG	1319
Db	1328	AAATCCGAGGATATTAACCTCGATGAGCCAAATGCAAGAGTGAATTACATATAGCAG	1387
Qy	1320	CTCTATGAAACAAATTCATGATGAGTGGCAAAATGCAAGAGTGAATTCCGAGTGCACA	1379
Db	1388	CTCTATGAAACAAATTCATGATGAGTGGCAAAATGCAAGAGTGAATTCCGAGTGCACA	1437
Qy	1380	GTTCGAGAGTTTCCACCTCGTGTCTCTGGATCTTGAACCTCTGAGCTGCGGCTC	1439
Db	1438	-----CAGAGTTTCCACACCGTTTCTTGGGCCCTTGAACCTCTGAGCTGCGGCTC	1492
Qy	1440	TGTCTTCTGAGGATCAACCACTTACTGAGCAAGAGAGAGAGTCCATCCCATGACAG	1499
Db	1493	TGTCTTCTGAGGATCAACCACTTACTGAGCAAGAGAGAGAGTCCATCCCATGACAG	1552
Qy	1500	AAGCAGACCGTTTTCAGCTCCAGTACTGCGGATTTGCCAAAAGCAAAACCCCGGCGC	1559
Db	1553	AAGCAGACCGTTTTCAGCTCCAGTACTGCGGATTTGCCAAAAGCAAAACCCCGGCGC	1612
Qy	1560	TGACTTGTGTGATCCCTCGATCCGCGAATGCGAGATAAATTAAGATTAATAAATTC	1619
Db	1613	TGACTTGTGTGATCCCTCGATCCGCGAATGCGAGATAAATTAAGATTAATAAATTC	1672
Qy	1620	TGACTTGTGAAATGCTTGTGGTGCATAAACCTTCAAGGAGACATCCAGACCGCTCA	1679
Db	1673	TGACTTGTGAAATGCTTGTGGTGCATAAACCTTCAAGGAGACATCCAGACCGCTCA	1732
Qy	1680	GTAACGCTCCATAGAGTTTAAATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAG	1739
Db	1733	GTATAGTCTCATAGAGTTTAAATAGGAGCAGGCTACAGCACCTTCCAGACATTTGGAG	1792
Qy	1740	CACGCGTGTATGCTTGTGAGCTGGCAACCGGAGATTTATTTGTTGAACACACATTTCTGG	1799
Db	1793	TACAGCTTGCATGGCAATTTAGCTCGCAACAGGAGATTTTGTTCGAAACCGCATTTCTGG	1852
Qy	1800	GGAAGACTTATTCAGAGACGAAAGCCACATAGCCCAATCATAGAGCTGTAGGAGTAT	1859
Db	1853	GGAAGACTTATTCAGAGATGAAGACCAATAGCCCAATCATAGAGCTGTAGGAGTAT	1912
Qy	1860	TCCAAGGCACTTTGTCTCTATCTGGAAATATTTCTCGGAAATTTCTTCAATCCAGAGGAGA	1919
Db	1913	CCCAAGGCACTTTGTCTCTGCTGCGAAATATTTCTCGGAAATTTCTTCAATCCAGAGGAGA	1972
Qy	1920	ACTCGACACATCACCAAGCTGAGCCCTCGAGCTCTTTGATGATCTTTGTTGAAAAAGTA	1979
Db	1973	ACTCGGACACATCACCAAGCTGAGCCCTCGAGCTCTTTGATGATCTTTGTTGAAAAAGTA	2032
Qy	1980	TGGCTGGCCCATGAAAGATGCTGCAAGTTTACAGATTTCTCTGATCCGATGTTTGAAT	2039
Db	2033	TGGCTGGCCCATGAAAGATGCTGCAAGTTTACAGATTTCTCTGATCCCAATGTTAGAT	2092
Qy	2040	GGTTCAGAAAAACAGGCTCAGCTGGGAAATGCGCTTGGCAATCTCTGTTGAATTTCT	2097
Db	2093	GGTTCAGAAAAACAGGCTCAGCTGGGAAATGCGCTTGGCAATCTCTGTTGAATTTCT	2150

RESULT 14
BC062941
LOCUS
DEFINITION
ACCESSION

BC062941 4287 bp mRNA linear ROD 11-DEC-2003
Mus musculus serine/arginine-rich protein specific kinase 2, mRNA
(cDNA clone IMAGE:6842723), complete cds.

This Page Blank (uspto)


```
QY 967 TTATTTGGAGAGCGCTCGAGAGATAGAGAAATTTGGAGCGGAGAGCTGAAGGAAATA 1026
Db 61 TTATTTGGAGAGCGCTCGAGAGATAGAGAAATTTGGAGCGGAGAGCTGAAGGAAATA 120
QY 1027 ATAGAAGAAACATCACTCAGCTGCACCTTCCAAATGACAGGATGGCGAATACTGCCCCA 1086
Db 121 ATAGAAGAAACATCACTCAGCTGCACCTTCCAAATGACAGGATGGCGAATACTGCCCCA 180
QY 1087 GAGGTGAAACTTAAACAAACAGGATTAGAGAGGCGGCTGAGGCGAGACTGCAAGGAC 1146
Db 181 GAGGTGAAACTTAAACAAACAGGATTAGAGAGGCGGCTGAGGCGAGACTGCAAGGAC 240
QY 1147 AATGCTGAGCTGAGGACGAGAGCAAGAAAGAGATGCTGAGAAAGAAAACATTGAAAAA 1206
Db 241 AATGCTGAGCTGAGGACGAGAGCAAGAAAGAGATGCTGAGAAAGAAAACATTGAAAAA 300
QY 1207 GATGAAGATGATGATAGATCAGGAACTTGGAAACATAGACCCCTACGTGGATAGAAATCACT 1266
Db 301 GATGAAGATGATGATAGATCAGGAACTTGGAAACATAGACCCCTACGTGGATAGAAATCACT 360
QY 1267 AAAACCAATGGCCATATTGAGAAATGGCCCATTTCTCACTGGAGCAGCAACTGGACCATGAA 1326
Db 361 AAAACCAATGGCCATATTGAGAAATGGCCCATTTCTCACTGGAGCAGCAACTGGACCATGAA 420
QY 1327 GATGATGATGAAGAGACTGCCCAATCTGAGGAATATAATCTTGTGATGAGCAAAATGCA 1386
Db 421 GATGATGATGAAGAGACTGCCCAATCTGAGGAATATAATCTTGTGATGAGCAAAATGCA 480
QY 1387 GAAAGTGATTACACATATAGAGCTCTATGAACAAATCAATGGTGAATTCGCAAAATGGA 1446
Db 481 GAAAGTGATTACACATATAGAGCTCTATGAACAAATCAATGGTGAATTCGCAAAATGGA 540
QY 1447 CGACATAAATTTCCGAGTCCAGATTTCCAGAGTTTCCACCTCGTGTCTCTCGATCC 1506
Db 541 CGACATAAATTTCCGAGTCCAGATTTCCAGAGTTTCCACCTCGTGTCTCTCGATCC 600
QY 1507 TTAGAACCCTGTGCGCTGCGGCTCTGTGCTTTCTGAGGATCACCACTTACTGAGCAAGAG 1566
Db 601 TTAGAACCCTGTGCGCTGCGGCTCTGTGCTTTCTGAGGATCACCACTTACTGAGCAAGAG 660
QY 1567 GAGAGCAGTCCATCCATGACAGAGACGAAACGTTTTAGCCTCAGCTCAGGATTTG 1626
Db 661 GAGAGCAGTCCATCCATGACAGAGACGAAACGTTTTAGCCTCAGCTCAGGATTTG 720
QY 1627 CAAAAGCAAAACCCGGCAGCTGACTTGTGTGAATCCCTGGATCCGCGGAATGCA 1686
Db 721 CAAAAGCAAAACCCGGCAGCTGACTTGTGTGAATCCCTGGATCCGCGGAATGCA 780
QY 1687 GATAAAATTAGAGTAAATAATTTGCTGACCTGGGAAATGCTTGTGGTGCATAAACACTTC 1746
Db 781 GATAAAATTAGAGTAAATAATTTGCTGACCTGGGAAATGCTTGTGGTGCATAAACACTTC 840
QY 1747 ACGGAAGACATCCAGACGCTCAGTACCGCTCCATAGAGGTTTTAATAGAGCGGGTAC 1806
Db 841 ACGGAAGACATCCAGACGCTCAGTACCGCTCCATAGAGGTTTTAATAGAGCGGGTAC 900
QY 1807 AGCACCCCTCGGACATCTGAGCAGCGCTGTATGTCATTTGAGCTGGCAACGCGGAGAT 1866
Db 901 AGCACCCCTCGGACATCTGAGCAGCGCTGTATGTCATTTGAGCTGGCAACGCGGAGAT 960
QY 1867 TATTTGTTTGAACCAATCTTGGGGAAGACTATTCCAGAGACGAAACCAATAGCCCCAC 1926
Db 961 TATTTGTTTGAACCAATCTTGGGGAAGACTATTCCAGAGACGAAACCAATAGCCCCAC 1020
QY 1927 ATCATAGAGCTGTAGGAGTATTCAGGCACTTGTCTATCTGGAAATATTTCTCGG 1986
Db 1021 ATCATAGAGCTGTAGGAGTATTCAGGCACTTGTCTATCTGGAAATATTTCTCGG 1080
QY 1987 GAATTTCTTCAATCGCAGAGAGAACTCGGACACATCACCAAGCTGAAGCCCTCGAGCCTC 2046
Db 1081 GAATTTCTTCAATCGCAGAGAGAACTCGGACACATCACCAAGCTGAAGCCCTCGAGCCTC 1140
```

```
QY 2047 TTTGATGATCTTGTGGAAGAGTATGCTGGCCCCCATGAAGATGCTGCACAGTTTACAGAT 2106
Db 1141 TTTGATGATCTTGTGGAAGAGTATGCTGGCCCCCATGAAGATGCTGCACAGTTTACAGAT 1200
QY 2107 TTCTGATCCGATGTTAGAAATGGTTCAGAAAGAAACGAGCCTCAGCTGGCGAATGCTT 2166
Db 1201 TTCTGATCCGATGTTAGAAATGGTTCAGAAAGAAACGAGCCTCAGCTGGCGAATGCTT 1260
QY 2167 CGGCATCCTTGTGTAATCTTAGCAAAATCTTACCAATATTGCAATCTGAGCTAGCAAAAT 2226
Db 1261 CGGCATCCTTGTGTAATCTTAGCAAAATCTTACCAATATTGCAATCTGAGCTAGCAAAAT 1320
QY 2227 GTTCCAGTACATTTGGACCTTAAACGGTGACTCTCATCTTTTAAACAGGATTAAGTGAGC 2286
Db 1321 GTTCCAGTACATTTGGACCTTAAACGGTGACTCTCATCTTTTAAACAGGATTAAGTGAGC 1380
QY 2287 TGGCTTCATCTCAGACCTTTATTTGCTTTGAGGTACTGTTGTTGACATTTTGTCTTTT 2346
Db 1381 TGGCTTCATCTCAGACCTTTATTTGCTTTGAGGTACTGTTGTTGACATTTTGTCTTTT 1440
QY 2347 TGTGCACTGTGATCCTGGGAAGGAGTACTTTTGTCTTTCAGCTAAGTAGTTTACTGACC 2406
Db 1441 TGTGCACTGTGATCCTGGGAAGGAGTACTTTTGTCTTTCAGCTAAGTAGTTTACTGACC 1500
QY 2407 ATTTTCTTCTGGAAACAAATAACATGCTCTAAGCAATGTTTCTGTGTTGTGACATTC 2466
Db 1501 ATTTTCTTCTGGAAACAAATAACATGCTCTAAGCAATGTTTCTGTGTTGTGACATTC 1560
QY 2467 AAATGTCAATTTTGTGATGAATAAATACTTTCCCTTTGTTTGGCAGGTTTGTGAAC 2526
Db 1561 AAATGTCAATTTTGTGATGAATAAATACTTTCCCTTTGTTTGGCAGGTTTGTGAAC 1620
QY 2527 TATTTATGAAGAAATATTTTAGCTGAGTACTATATAAATTTACAATCTTAGAAATTTATCA 2586
Db 1621 TATTTATGAAGAAATATTTTAGCTGAGTACTATATAAATTTACAATCTTAGAAATTTATCA 1680
QY 2587 AGTTGGGACCAAGAAATAGCAAGGAAATGTACAATTTATCTCTGCGCAAGGAGCA 2646
Db 1681 AGTTGGGACCAAGAAATAGCAAGGAAATGTACAATTTATCTCTGCGCAAGGAGCA 1737
QY 2647 TCATTCTCTGATATTAGTGTATGTAAT 2674
Db 1738 TCATTCTCTGATATTAGTGTATGTAAT 1765
```

```
RESULT 2
BX463511
LOCUS
DEFINITION BX463511 Homo sapiens ADULT BRAIN Homo sapiens cdna clone
CS0DN003YA14 5-PRIME, mRNA sequence.
BX463511
BX463511.2 GI:47056867
EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 995)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT
On May 22, 2003 this sequence version replaced gi:31027556.
Contact: Genoscope
Genoscope - Centre National de Sequencage
Bp 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4027.r
For more information about this cluster, see
```

http://www.genoscope.cns.fr/cdna?b=CS0DN003BA07QP1&c=4027.r.

FEATURES
source
Location/Qualifiers
1..995
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DN003YAL14"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match 30.0%; Score 975.6; DB 5; Length 995;
Best Local Similarity 99.4%; Pred. No. 6.2e-209;
Matches 986; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
QY 910 AAAATATCTAAACAAAGAAAGAAACCTGAAAGAAACAGAGAGCGAGCTGAGTTA 969
DB 5 AAAATATCTAAACAAAGAAAGAAACCTGAAAGAAACAGAGAGCGAGCTGAGTTA 64
QY 970 TTGGAGAGCGCTGCAGAGATAGAGAAATTGGAGCGAGAGAGCTGAAAGGAAATATA 1029
DB 65 TTGGAGAGCGCTGCAGAGATAGAGAAATTGGAGCGAGAGAGCTGAAAGGAAATATA 124
QY 1030 GAAGAAACATCACTCAGCTGACCTTCCATATGACAGAGTGGCGAATCTGCCAGAG 1089
DB 125 GAAGAAACATCACTCAGCTGACCTTCCATATGACAGAGTGGCGAATCTGCCAGAG 184
QY 1090 GTGAAACTTAAACAAACAGATTAGAGAGCGGCTGAGCGAGAGCTCCAAAGGACAT 1149
DB 185 GTGAAACTTAAACAAACAGATTAGAGAGCGGCTGAGCGAGAGCTCCAAAGGACAT 244
QY 1150 GTGAGCTGAGGACAGAGAGAGAGAGATGCTGAGAAAGAAACATTTGAAAGAT 1209
DB 245 GTGAGCTGAGGACAGAGAGAGAGAGATGCTGAGAAAGAAACATTTGAAAGAT 304
QY 1210 GAAGATGATTAGATCAGGAACTTGGGAAACATAGACCTTACCTGGATAGAAATCACTAA 1269
DB 305 GAAGATGATTAGATCAGGAACTTGGGAAACATAGACCTTACCTGGATAGAAATCACTAA 364
QY 1270 ACCAATGGCCATATTGAGAAATGCCCATTTCTCACTGGAGCAGCAACTGGACGATGAAGAT 1329
DB 365 ACCAATGGCCATATTGAGAAATGCCCATTTCTCACTGGAGCAGCAACTGGACGATGAAGAT 424
QY 1330 GATGATGAAGAGACTGCCCAATCTTGAGGAATATAATCTTGATGAGCAGCAAAATGCAGAA 1389
DB 425 GATGATGAAGAGACTGCCCAATCTTGAGGAATATAATCTTGATGAGCAGCAAAATGCAGAA 484
QY 1390 AGTGATTACATATAGCAGCTCTTATGAACAAATCAATGGTGAATTTGCAAAATGCAGCA 1449
DB 485 AGTGATTACATATAGCAGCTCTTATGAACAAATCAATGGTGAATTTGCAAAATGCAGCA 544
QY 1450 CATAAATTCGGAGTCAAGTTTCCAGAGTTTCCACCTCGTGTCTCTGGATCTTTA 1509
DB 545 CATAAATTCGGAGTCAAGTTTCCAGAGTTTCCACCTCGTGTCTCTGGATCTTTA 604
QY 1510 GAACTGTGGCTGCGGCTCTGTGCTTCTGAGGATACCACTTACTGAGCAAGAGGAG 1569
DB 605 GAACTGTGGCTGCGGCTCTGTGCTTCTGAGGATACCACTTACTGAGCAAGAGGAG 664
QY 1570 AGCAGTCCATCCATGACAGAGCAGAACCGTTTTCAGCTTCCAGTCTCGGGATTTGCCA 1629
DB 665 AGCAGTCCATCCATGACAGAGCAGAACCGTTTTCAGCTTCCAGTCTCGGGATTTGCCA 724
QY 1630 AAAGCAAAACCCGGGACGCTGACTTGTGTTGGTGAATCCCTCGGATCCGGGAAATGCAGAT 1689
DB 725 AAAGCAAAACCCGGGACGCTGACTTGTGTTGGTGAATCCCTCGGATCCGGGAAATGCAGAT 784

QY 1690 AAAATTTAGAGTAAAAATTGCTGACCTGGGAAATGCTTGTGGTGCATAAAACACTTCACG 1749
DB 785 AAAATTTAGAGTAAAAATTGCTGACCTGGGAAATGCTTGTGGTGCATAAAACACTTCACG 844
QY 1750 GAAGACATCCAGACGCGTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGC 1809
DB 845 GAAGACATCCAGACGCGTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGC 904
QY 1810 ACCCTGTGGGACATCTGGAGCAGCGGCTGTATGGCATTTGAGCTGCGCAACGGGAGATTAT 1869
DB 905 ACCCTGTGGGACATCTGGAGCAGCGGCTGTATGGCATTTGAGCTGCGCAACGGGAGATTAT 964
QY 1870 TTGTTTGAACCACTTCTGGGGAAGACATATTC 1901
DB 965 TTGTTTGAACCACTTCTGGGGAAGACATATTC 995

RESULT 3

BM464185
LOCUS AGENCOURT_6438899 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5535256
DEFINITION 5', mRNA sequence.
ACCESSION BM464185
VERSION BM464185.1 GI:18513227
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1108)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12223 row: a column: 17
High quality sequence stop: 787.

FEATURES

source
Location/Qualifiers
1..1108
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5535256"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2.1 Kb."

ORIGIN

Query Match 28.8%; Score 937.8; DB 4; Length 1108;
Best Local Similarity 94.7%; Pred. No. 2.2e-200;
Matches 1027; Conservative 0; Mismatches 47; Indels 11; Gaps 5;
QY 1367 ATCTTGATGAGCAAAATGCAGAAAGTGATTAACATATAGCAGCTCCTATGAACAATTC 1426
DB 21 ATCTTGATGAGCAAAATGCAGAAAGTGATTAACATATAGCAGCTCCTATGAACAATTC 80
QY 1427 ATGGTGAATTGCAATGACGACATAAATTTCCGAGTCAAGTCCAGAGTTTCCCA 1486
DB 81 ATGGTGAATTGCAATGACGACATAAATTTCCGAGTCAAGTCCAGAGTTTCCCA 140
QY 1487 CCTCGTTGTTCTCTGGATCCTTAGAACCTGTGGCTTCTGTGCTTCTTGAGGGAT 1546
DB 141 CCTCGTTGTTCTCTGGATCCTTAGAACCTGTGGCTTCTGTGCTTCTTGAGGGAT 200

QY	1547	CACCACTTACTGAGCAAGAGGAGGAGCTGCTCATCCATGACAGAGCAGACGGTTTCAG	1606
Db	201	CACCACTTACTGAGCAAGAGGAGGAGCTGCTCATCCATGACAGAGCAGACGGTTTCAG	260
QY	1607	CCTCAGTACTGGGATTTGCAAAAGCAAAACCCGGGAGCTGACTTTGTTGGTGAATC	1666
Db	261	CCTCAGTACTGGGATTTGCAAAAGCAAAACCCGGGAGCTGACTTTGTTGGTGAATC	320
QY	1667	CCTCGATCCGGGATTCAGATAAAATTAGAGTAAATAATTGCTGACCTGGGAATGCTT	1726
Db	321	CCTCGATCCGGGATTCAGATAAAATTAGAGTAAATAATTGCTGACCTGGGAATGCTT	380
QY	1727	GTTGGGTGATAAACACTTCCAGGAAGATCCAGACGGCTCAGTACCGTCCATAGAGG	1786
Db	381	GTTGGGTGATAAACACTTCCAGGAAGATCCAGACGGCTCAGTACCGTCCATAGAGG	440
QY	1787	TTTTTAATAGAGCGGGTACAGCACCCCTCGGACATCTGGAGCAGCGCTGTATGGCAT	1846
Db	441	TTTTTAATAGAGCGGGTACAGCACCCCTCGGACATCTGGAGCAGCGCTGTATGGCAT	500
QY	1847	TTGAGCTGCAACCGGAGATTAATTTGTTGAACCACTCTCTGGGAAGACTATTCACAG	1906
Db	501	TTGAGCTGCAACCGGAGATTAATTTGTTGAACCACTCTCTGGGAAGACTATTCACAG	560
QY	1907	ACGAAGACCACATAGCCACATAGAGCTGCTAGGAGTATTCAGAGGCACTTTGCTC	1966
Db	561	ACGAAGACCACATAGCCACATAGAGCTGCTAGGAGTATTCAGAGGCACTTTGCTC	620
QY	1967	TATCTGGAATAATTTCTCGGGAATTTCTCAATCGCAGAGGAACTGGGACACATCACA	2026
Db	621	TATCTGGAATAATTTCTCGGGAATTTCTCAATCGCAGAGGAACTGGGACACATCACA	680
QY	2027	AGCTGAAGCCCTGGAGCCTTTGATGATCTTTGTAAGATGATGCTGGCCCATGAG	2086
Db	681	AGCTGAAGCCCTGGAGCCTTTGATGATCTTTGTAAGATGATGCTGGCCCATGAG	740
QY	2087	ATGCTGCACAGTTTACAGATTTCTGATCCCGATGTTAGAAATGTTTCCAGAAAAACGAG	2146
Db	741	ATGCTGCACAGTTTACAGATTTCTGATCCCGATGTTAGAAATGTTTCCAGAAAAACGAG	800
QY	2147	CCTCAGTGGCAATGCTTGGGCACTTCTGTTGAATTTCTAGCAATTTCTACCAATAT	2206
Db	801	CCTCAGTGGCAATGCTTGGGCACTTCTGTTGAATTTCTAGCAATTTCTACCAATAT	860
QY	2207	TGCAATCTGAGCTAGCAAAATGTTCCAGTACATT - GGACTAAACGGTGACTCTCATCT	2265
Db	861	TGCAATCTGAGCTAGCAAAATGTTCCAGTACATTGTTGGACCTTAACCGTGACTCTCATCT	920
QY	2266	TTAACAGGATTACAAGTGAAGCTGGCTTCATCTCAGACCTTTATTTTGTGTTGAGTACT	2325
Db	921	TTTACAGGATTACAAGTGAAGCTGGCTTC - TCCTCAGACCTTAATTTTGTGTTTAAAGGACT	979
QY	2326	GTTGTTGACATTTGCTTTTGTGCACTGTA - - - - - TCCTGGGAAGGTTAGTCTTTT	2380
Db	980	GTTGTTTGTACATTTGCTTTTGTGCTGTGGAATCCCTGGGGAAGGTTAACCTTTT	1039
QY	2381	GTCTTTCAGCTA - AGTAGTTTACTGACCAATTTCTTCT - - - GGAAACAATAACATGCTCT	2436
Db	1040	GGCTCTGAGCTAGAGTTTACTGACCAATTTTCTTCTGGAACAATAACATGCTCT	1099
QY	2437	AAGCA 2441	
Db	1100	TAACA 1104	
RESULT 4	AK035962	LOCUS	AK035962
DEFINITION	Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630021E16 product:serine/arginine-rich protein specific kinase 2, full insert sequence.		
ACCESSION	AK035962		

VERSION	AK035962.1	GI:26331059
KEYWORDS	HTC; CAP trapper.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
AUTHORS	Carninci, P. and Hayashizaki, Y.	
TITLE	High-efficiency full-length cDNA cloning	
JOURNAL	Weth. Enzymol. 303, 19-44 (1999)	
MEDLINE	99279253	
PUBMED	10349636	
REFERENCE	1	
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.	
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes	
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)	
MEDLINE	20499374	
PUBMED	11042159	
REFERENCE	3	
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Hara, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.	
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer	
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)	
MEDLINE	20530913	
PUBMED	11076861	
REFERENCE	4	
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.	
TITLE	Functional annotation of a full-length mouse cDNA collection	
JOURNAL	Nature 409, 685-690 (2001)	
MEDLINE		
PUBMED		
REFERENCE	5	
AUTHORS	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.	
TITLE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs	
JOURNAL	Nature 420, 563-573 (2002)	
MEDLINE		
PUBMED		
REFERENCE	6 (bases 1 to 3653)	
AUTHORS	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, I., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohata, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, I., Sobabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (16-JUN-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)	
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/.	

FEATURES		Location/Qualifiers	
Source		1..3653	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/strain="C57BL/6J"	
		/db_xref="FANTOM DB:9630021E16"	
		/db_xref="taxon:10090"	
		/clone="9630021E16"	
		/issue_type="cerebellum"	
		/clone_lib="RIXEN full-length enriched mouse cDNA library"	
		/dev_stage="16 days neonate"	
CDS		<1..667	
		/note="unnamed protein product; putative serine/arginine-rich protein specific kinase 2 (MGDI) [GI:1201408, GB NM_009274, evidence: BLASTN, 99%, match=921]"	
		/codon_start=2	
		/protein_id="BAC29260.1"	
		/db_xref="GI:26331060"	
		/translation="GSVISEGSLTEQESSPSHDSRTVSASSSTDLPKTKYRAADL LYNLDPRNADKIRKVIADLGNACVWHKHFTIEDIQTRYSIEVLIGAGYSTPADLWS TACMAPEADLYLFEPHSGEDYSRDEHIAHIIELLSIPRHFALSGKYSREFFNRR GELRHITKLPWSLFDVLVEKYGPHEDAQAQFTDFLIPMLEMVPKRSAGSCLRHWP LNS"	
polyA_signal		3627..3632	
		/note=".putative"	
polyA_site		3653	
		/note=".putative"	
ORIGIN			
		Query Match 28.8%; Score 936.4; DB 3; Length 3653;	
		Best Local Similarity 91.5%; Pred. No. 5.7e-200;	
		Matches 1049; Conservative 0; Mismatches 86; Indels 11; Gaps 5;	
QY	1525	GGCTCTGTCTTCTGAGGGATCCACCTTACTTGAGCAAGAGGAGAGCAGTCCATCCCAT	1584
DB	2	GGCTCTGTGATTTACAGGGATCGCCACTTACCGACGAGGAGGAAGCAGTCCCTCCCAT	61
QY	1585	GACAGAGCAGAACGGTTTCAGCTTCAGTCTAGTCTGGGGATTTGCCAAAAGCAAAACCCGG	1644
DB	62	GACAGAGCAGGACAGTTTCAGCTCTAGTCTAGTCTGGAGATTTGCCAAAACCAAAACCCGG	121
QY	1645	GCAGCTGACTTCTGGTGAATCCCTGGATCCCGGAATGCAGATTAATAATTAGATATA	1704
DB	122	CGGGCTGACCTGTGGTGAACCTCTGGATCCACGGAATGCAGATTAATAATTAGATATA	181
QY	1705	ATTGCTGACTCGGAAATCTTGTGGTGCAATAACACTTCACGGAAGACATCCAGAGC	1764
DB	182	ATTGCTGACTCGGAAATCTTGTGGTGCAATAACATTTACAGAGGATATCCAGACA	241
QY	1765	CGTCACTACCGTCCATAGAGTTTAAATAGAGCGGGGTACAGCACCCCTCGGACATC	1824
DB	242	CGTCACTATAGTCCATAGAGTTTAAATAGAGCGAGGCTACAGCACACCTCGACAT	301
QY	1825	TGAGCAGCGCTGTATGCAATTTGAGCTGGCAACGGGAGATATTGTTGAACCAT	1884
DB	302	TGGAGTACAGCTTGCATGCAATTTGAGCTCGCCACAGGAGCTATTGTTGCAACCGCAT	361
QY	1885	TCTGGGGAAGACTATTTCCAGAGCAGGAAGACCATATAGCCCATCATATAGAGTGTAGGC	1944
DB	362	TCTGGGGAAGACTATTTCCAGAGATGAGACCATATAGCCCATCATATAGAGTGTAGGC	421
QY	1945	AGTATTTCCAGGACATTTCTCTATCTGGAATAATATTTCTGGGAATTTCTCAATCGCAGA	2004
DB	422	AGTATTTCCAGGACATTTCTCTCTGCTGGAATAATATTTCTGGGAATTTCTCAATCGCAGA	481
QY	2005	GGAGAACTCGGACATCATCACCAGCTGAAGCCCTGGAGCTCTTGTATGATCTTGTGGAA	2064
DB	482	GGAGAACTCGGACATCATCACCAGCTGAAGCCCTGGAGCTCTTGTATGATCTTGTGGAA	541
QY	2065	AGATATGGCTGGCCCCCATGAAGATGCTGCACAGTTTACAGATTTCTCTGATCCCGATGTTA	2124
DB	542	AAGATATGGCTGGCCCCCATGAAGATGCTGCACAAATTTACAGATTTCTCTGATCCCAATGTTA	601
QY	2125	GAATGGTTCCAGAAAAACGAGCTCAGCTGCGGAATGCTTCCGGCATCTCTGGTTGAAT	2184
DB	602	GAGATGGTTCCAGAAAAACGAGCTCAGCTGCGGAATGCTTCCGACATCTCTGGTTGAAT	661
QY	2185	TCTTAGCAAAATCTACCAATATTGCAATCTGAGCTAGCAAAATGT-TCCCAGTACATTGGA	2243
DB	662	TCTTAGCAATTTCTACAAATATAGCATTTCTGAGCTAGCAAAATTTCTCCAGTACATTGGA	721
QY	2244	CCTAACCGGTGACTCTCAATCTTTTAAACAGGATTACAGGTGAGCTGGCTTCATCTCAGAC	2303
DB	722	CCTACACAGTGGCTCTCAATCTTTTAAACAGGATTACAGGTGAGCTGGCTCCACCTCAGAC	781
QY	2304	CTTTATTTTGGCTTTCAGGTACTGTTGTTTGAATTTTGTGCTTTTGTGCACTGTGATCTCG	2363
DB	782	CTTTGTAGTGGCTTTCAGGTACTGTTGTTTGAATTTTGTGCTTTTGTGCACTGTGATCTCG	841
QY	2364	GGGAAGGTGAGTCTTTTGTCTTTCAGCTAGTAGTATTACTGACATTTTCTCTCGGAACA	2423
DB	842	GGGAAGGTGGTCT-TTGTCTTTCAGTAAAGTAGTATTACTGGCAATTTTCTCTCGGAACA	900
QY	2424	ATAACATGTCTCTAAGCAATTTTCTTGTGTTGTGTGACATTTCAAAATGTC-----ATTTT	2478
DB	901	ATAACATGTCTCTAAGCAATTTTCTTGTGTTGTGTGACATTTCAAAATGTC-----ATTTT	960
QY	2479	TTTGAATGAAAATACCTTTCCCTTGTGTTTGGCAGGTTTGTGCACTATTATATGAAGA	2538
DB	961	TTTGAACGAAAATACCTTTCCCTTGTGTTTGGCAGGTTTGTGCACTATTATATGAAGA	1020
QY	2539	ATATTTTAGCTGAGTACTATATAATTTTACAATCTTAAGAAATTTACAAGTTGGGAACCA	2598
DB	1021	CATATTTTAGCTGAGTATTATATAATTTTACAATCTTAAGAAATTTATCA--GTTGGAAACA	1078
QY	2599	AGAAATAGCAAGGAAATGTACAAATTTTATCTTCTGCAAAAGGACATCATTTCTGTAT	2658
DB	1079	AGAAA--GAGCAAGGAAATGTACAAATTTTATCTGTCGCAAGGGAATCATTTCTGTAT	1136
QY	2659	TATAGT 2664	
DB	1137	TATACT 1142	
RESULT 5			
BX384358		1001 bp mRNA linear EST 26-APR-2004	
BX384358		Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens	
DEFINITION		CDNA clone CS0DK009YCL1 5-PRIME, mRNA sequence.	
ACCESSION		BX384358	
VERSION		BX384358.2 GI:46574523	
KEYWORDS		EST.	
SOURCE		Homo sapiens (human)	
ORGANISM		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE		1. (bases 1 to 1001)	
AUTHORS		Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
TITLE		Full-length cDNA libraries and normalization	
JOURNAL		Unpublished (2001)	
COMMENT		On May 8, 2003 this sequence version replaced gi:30451233.	
		Contact: Genoscope	
		Genoscope - Centre National de Sequencage	
		BP 191 91006 EVRY cedex - France	
		Email: seqrefgenoscope.cns.fr, Web : www.genoscope.cns.fr	
		1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime	
		end enriched, double-strand cDNA was digested with Not I and cloned	
		into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library	
		was normalized. Library was constructed by life technologies, a	
		division of Invitrogen. This sequence belongs to sequence cluster	
		4027.r	
		For more information about this cluster, see	
		http://www.genoscope.cns.fr/cdna?s=CS0DK009AB06QP1&c=4027.r.	
FEATURES		Location/Qualifiers	
source		1..1001	

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DK009VC11"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/cell_line="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoRV sites of the Not I and EcoRV
sites of the pCMVSPORT 6 vector. Library was normalized."
```

ORIGIN

Query Match	28.5%;	Score	927.6;	DB	5;	Length	1001;
Best Local Similarity	97.8%;	Pred. No.	4.2e-198;				
Matches	960;	Conservative	10;	Mismatches	9;	Indels	3;

3; Gaps 3;

QY	613	ATGGTCTTCGAAGTACTTGGCCACCATCTCCCAAGTGGATCATCAAAATCAACTATCAAA	672
DB	1	ATGGTCTTCGAAGTACTTGGCCACCATCTCCCAAGTGGATCATCAAAATCAACTATCAAA	60
QY	673	GGCTCCCAAGTACTTGGTGAAGATATCATTCGACAGGTCCTTCAAGGGTTAGATTAC	732
DB	61	GGCTCCCAAGTACTTGGTGAAGATATCATTCGACAGGTCCTTCAAGGGTTAGATTAC	120
QY	733	TTACACAGTAAGTGCAGATCATCTCATCTGCATATAAGCCGGAAATATCTTGATGTGT	792
DB	121	TTACACAGTAAGTGCAGATCATCTCATCTGCATATAAGCCGGAAATATCTTGATGTGT	180
QY	793	GTGGATGATGCATATCTGAGAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGT	852
DB	181	GTGGATGATGCATATCTGAGAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGT	240
QY	853	GCTCTCTCTTCAGGCTTCAGTGCAGTACGGCTCCACAGCAGAACTATAGAAAA	912
DB	241	GCTCTCTCTTCAGGCTTCAGTGCAGTACGGCTCCACAGCAGAACTATAGAAAA	300
QY	913	ATATCTAAACAAAGAAAGAAATCTGAAAGAAAGAAAGAAAGAAAGAAAGAAATGTTG	972
DB	301	ATATCTAAACAAAGAAAGAAATCTGAAAGAAAGAAAGAAAGAAAGAAAGAAATGTTG	360
QY	973	GAGAAGCGCTCGAGGAGATAGAAAGATTGGACGAGAGCTGAAAGAAAGAAATATAGAA	1032
DB	361	GAGAAGCGCTCGAGGAGATAGAAAGATTGGACGAGAGCTGAAAGAAAGAAATATAGAA	420
QY	1033	GAAACATCATCTCAGCTCAGCTTCCATATGACAGATGCGCAATACTGCCAGAGGTG	1092
DB	421	GAAACATCATCTCAGCTCAGCTTCCATATGACAGATGCGCAATACTGCCAGAGGTG	480
QY	1093	AAACTAAACAAACAGGATAGAGGCGGCTGAGCGAGAGCTGCAAGGGAATGGT	1152
DB	481	AAACTAAACAAACAGGATAGAGGCGGCTGAGCGAGAGCTGCAAGGGAATGGT	540
QY	1153	GAAGCTGAGGACACGGAAGAAAGAGATGCTGAGAAAGAAACATTTGAAAGATGAA	1212
DB	541	GAAGCTGAGGACACGGAAGAAAGAGATGCTGAGAAAGAAACATTTGAAAGATGAA	600
QY	1213	GATGATGTAGATCAGGAATCTGCAACATAGACCTACCTGAGTGAATCACCTAAAC	1272
DB	601	GATGATGTAGATCAGGAATCTGCAACATAGACCTACCTGAGTGAATCACCTAAAC	660
QY	1273	AATGSCCATATTGAGATGGCCCATTTCTACTGAGAGAGCACTGGACGATGAAGATGAT	1332
DB	661	AATGSCCATATTGAGATGGCCCATTTCTACTGAGAGAGCACTGGACGATGAAGATGAT	720
QY	1333	GATGAGAGAGCTGCCAAATCTGAGGAATATATCTTGATGAGCCAAATGCGAAAGT	1392
DB	721	GATGAGAGAGCTGCCAAATCTGAGGAATATATCTTGATGAGCCAAATGCGAAAGT	780
QY	1393	GATTACACATATAGCAGCTCTCTATGAACAAATTCATATGGTGAATTCGCAATGACGACAT	1452
DB	781	GATTACACATATAGCAGCTCTCTATGAACAAATTCATATGGTGAATTCGCAATGACGACAT	840

QY	1453	AAAAATCCCGAGTCCACAGTCCACAGTTCCTCCACCTGTTGTTCTCTGGATCCTTAGAA	1512
DB	841	AAAAATCCCGAGTCCACAGTTCCTCCACCTGTTGTTCTCTGGATCCTTAGAA	900
QY	1513	CTGTGGCTTCGGGTCTGTGCTTTCTGAGGATCACCACCTTACTGAGCAGAGGAGAGC	1572
DB	901	CTGTGGCTTCGGGTCTGTGCTTTCTGAGGATCACCACCTTACTGAGCAGAGGAGAGC	957
QY	1573	AGTCCATCCATGACAGAGCA	1594
DB	958	AGTCCATCCGACAGAGCA	979

RESULT 6

BX463510/c 1025 bp mRNA linear EST 05-MAY-2004

LOCUS BX463510 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone

DEFINITION CS0DN003YA14 3-PRIME, mRNA sequence.

ACCESSION BX463510

VERSION BX463510.2 GI:47055934

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1025)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On May 22, 2003 this sequence version replaced gi:31025523.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 4027.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?s=CS0DN003BA07NP1&c=4027.r.

FEATURES

Location/Qualifiers

1..1025

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DN003YA14"

/tissue_type="ADULT BRAIN"

/dev_stage="adult"

/clone_lib="Homo sapiens ADULT BRAIN"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

Query Match	27.9%;	Score	909.2;	DB	5;	Length	1025;
Best Local Similarity	94.0%;	Pred. No.	6e-194;				
Matches	972;	Conservative	7;	Mismatches	46;	Indels	9;

4; Gaps 4;

QY	1641	CCGGCAGCTGACTTGTGTGTTGAATCCCTCGATCCCGGAATCGAGATAAATAGAGT	1700
DB	1025	CCGGCAGCTGACTTGTGTGTTGAATCCCGGAATCGAGATAAATAGAGT	968
QY	1701	AAAAATTCGTGACCTGGGAATGCTTGGTGTGATTAACACTTCACGGAAGACATCA	1760
DB	967	AAAAATTCGTGACCTGGGAATGCTTGGTGTGATTAACACTTCACGGAAGACATCA	908
QY	1761	GACGCTCAGTACCGCTCCATAGAGTTTATAGAGCGGGGTACAGACCCCTCGCGA	1820
DB	907	GACGCTCAGTACCGCTCCATAGAGTTTATAGAGCGGGGTACAGACCCCTCGCGA	848

FEATURES	Location/Qualifiers	
	1..1486	
source	/organism="Mus musculus"	
	/mol_type="mRNA"	
	/strain="C57BL/6J"	
	/db_xref="FANTOM DB:A630008K11"	
	/db_xref="taxon:10090"	
	/clone="A630008K11"	
	/tissue type="thymus"	
	/clone_lib="RIKEN full-length enriched mouse cDNA library"	
CDS	28..591	
	/note="unnamed protein product: putative serine/arginine-rich protein specific kinase 2 (MGD) [GI:1201408, GB NM_009274, evidence: BLASTN, 99%, match=921]"	
	/protein_id="BAC30942.1"	
	/db_xref="GI:26334483"	
	/translation="WAATRAADLLVNPDPNADKIRVKIADLGNACVHKHFTEDI QTRQSRLEVLIGAGTSPADIWSTACNAFELATGDYLFEPHSGEDYSRDRDHIHII ELIGSPHRFALSGYSREFFNRGELRHITKLPWSLFDVLVEKIGMHPHEDAAQFTD FLIPMLEVMPEKRSAGBCLRHPWLNS"	
ORIGIN		
	Query Match 26.1%; Score 847.6; DB 3; Length 1486;	
	Best Local Similarity 91.7%; Pred. No. S.1e-180;	
	Matches 954; Conservative 0; Mismatches 74; Indels 12; Gaps 5;	
QY	1632	AGCAAAACCCGGCAGCTGACTTGTGTGTAATCCCTCGATCCGCGAATGCGATAA 1691
DB	33	AGCAAAACCCGGCGCTGACTTGTGTGTAATCCCTCGATCCGCGAATGCGATAA 92
QY	1692	AATTAGATAAATTTGCTGACCTGGGAAATGCTTGTGGTGCATAAATGCTTCA 1751
DB	93	AATTAGATAAATTTGCTGACCTGGGAAATGCTTGTGGTGCATAAATGCTTCA 152
QY	1752	AGACATCCAGACGCTGACGACGCTCCATAGAGGTTTAAATAGAGCGGGGTACAG 1811
DB	153	GGATATCCAGACGCTGACGACGCTCCATAGAGGTTTAAATAGAGCGGGGTACAG 212
QY	1812	CCCTGGCGGACATCTGGAGCAGCGCTGTATGTCATTTGAGCTGGGAAATTTCT 1871
DB	213	ACCTGCGAGACATTTGGAGTACAGCTTGTGATGCTTGGCTGCGCCACAGAGACT 272
QY	1872	GTTTGAACACATTTCTGGGGAACATTTCCAGAGCAGAGACCATAGCCCATCAT 1931
DB	273	GTTTGAACACATTTCTGGGGAACATTTCCAGAGCAGAGACCATAGCCCATCAT 332
QY	1932	AGAGCTGCTAGGAGATTTCCAAAGGACATTTGCTCTATCTGGAATAATTTCTGG 1991
DB	333	AGAGCTGCTAGGAGATTTCCAAAGGACATTTGCTCTGCTGGAATAATTTCTGG 392
QY	1992	CTTCAATCGCAGAGGAGAACTGGGACACATCATCAAGCTGAAGCCCTGGAGCTT 2051
DB	393	CTTCAATCGCAGAGGAGAACTGGGACACATCATCAAGCTGAAGCCCTGGAGCTT 452
QY	2052	TGTACTTGTGGAAGATGCTGGGCGCCCATGAGATGCTGACAGTTTACAGATTCT 2111
DB	453	TGTACTTGTGGAAGATGCTGGGCGCCCATGAGATGCTGACAGTTTACAGATTCT 512
QY	2112	GATCCCGATGTTAGAAATGTTCCAGAAAAACAGAGCTCAGCTGGGGAATGCTTC 2171
DB	513	GATCCCGATGTTAGAAATGTTCCAGAAAAACAGAGCTCAGCTGGGGAATGCTTC 572
QY	2172	TCCTTGGTGAATTTCTAGCAATTTACCAATATGCAATTTGAGTGAATGCTTC 2230
DB	573	TCCTTGGTGAATTTCTAGCAATTTACCAATATGCAATTTGAGTGAATGCTTC 632
QY	2231	CCAGTACATTTGAGCTAAACGGTGAATCTCATTTTAAACAGGATTAACAGTGG 2290
DB	633	CCAGTACATTTGAGCTAAACGGTGAATCTCATTTTAAACAGGATTAACAGTGG 692

QY	2291	TTCATCTCAGACCTTTATTTTCTTTGAGTACTGTTTGTGACATTTTGTCTTTGTG 2350
DB	693	TCACACCTCAGACCTTTTGTAGTCTTTGAGTACTGTTTGTGACATTTTGTCTTTGTG 752
QY	2351	CACCTGATCCTCGGGAAGGTTAGTCTTTTGTCTTCTGAGTACTGTTTGTGACATTTT 2410
DB	753	CACCATGCTCTCGGGAAGGCTGGTC-TTTGTCTTCTGAGTACTGTTTGTGACATTTT 811
QY	2411	TCCTTGGAAACAAATACATGCTCTTAAGCAATTTTCTTGTGTTGTGTTGACATTTCA 2470
DB	812	TCCTTGGAAACAAATACATGCTCTTAAGCAATTTTCTTGTGTTGTGTTGACATTTCA 871
QY	2471	GTC-----ATTTTGTGAATGAAATATCTTCCCTTTTGTGTTTGTGAGGTTTGTGA 2524
DB	872	GTCCTTTTGTGTTTGTGAAAGAAATATCTTCCCTTTTGTGTTTGTGAGGTTTGTGA 931
QY	2525	ACTATTTATGAAGAAATATTTTGTGTTGAGTACTATATTTTACAATCTTAAAGAAAT 2584
DB	932	ACTATTTATGAAGAAATATTTTGTGTTGAGTACTATATTTTACAATCTTAAAGAAAT 991
QY	2585	CAAGTTGGGAACCAAGAAATAGCAAGGAAATGTAATTTTATCTTCTGCGAAAGGGA 2644
DB	992	CA--GTTGGAAACCAAGAA--GAGCAAGGAAATGTAATTTTATCTGCTGCGAAAGGGA 1047
QY	2645	CATCATTCCTGTATTATAGT 2664
DB	1048	AATCATTCCTGTATTATAGT 1067

RESULT 8

BX423897/c

LOCUS

DEFINITION BX423897 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone

ACCESSION BX423897

VERSION BX423897.2

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1014)

AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On May 13, 2003 this sequence version replaced gi:30638759.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 4027.r

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna/s=C51AN001ZE08NP1&c=4027.r.

Location/Qualifiers

1..1014

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CSODN003YA14"

/tissue type="ADULT BRAIN"

/dev stage="adult"

/clone_lib="Homo sapiens ADULT BRAIN"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

FEATURES

source

ORIGIN

[illegible]

```
Db 476 CAATCGCAGAGAGAAGCTGCGACACATCACCAAGCTGAAGCCCTGGAGCCCTCTTTGATGT 417
QY 2055 ACTTGTGGAAGATAGTGGCTGCCCCCATGAAGATGCTGCACAGTTTACAGATTTCTCTGAT 2114
Db 416 ACTTGTGGAAGATAGTGGCTGCCCCCATGAAGATGCTGCACAGTTTACAGATTTCTCTGAT 357
QY 2115 CCCGATGTTAGAAATGGTTTCCAGAAAAACGAGCCCTCAGCTGCGGAATGCCCTTCGGCATCC 2174
Db 356 CCCGATGTTAGAAATGGTTTCCAGAAAAACGAGCCCTCAGCTGCGGAATGCCCTTCGGCATCC 297
QY 2175 TTGGTTGAATCTTAGCAAAATCTACCAATATTTGCAATCTGAGCTAGCAAAATGTTCCGAG 2234
Db 296 TTGGTTGAATCTTAGCAAAATCTACCAATATTTGCAATCTGAGCTAGCAAAATGTTCCGAG 237
QY 2235 TACATTGGACCTTAACGGTGACTCTCATCTTTTAAACAGGATTAACAGTGAAGTGGCTTCA 2294
Db 236 TACATTGGACCTTAACGGTGACTCTCATCTTTTAAACAGGATTAACAGTGAAGTGGCTTCA 177
QY 2295 TCTCAGACCTTTATTTTCTTTGAGGTAAGTCTGTTTGAACATTTTGTGCACT 2354
Db 176 TCTCAGACCTTTATTTTCTTTGAGGTAAGTCTGTTTGAACATTTTGTGCACT 117
QY 2355 GTGATCTCGGAAGGATGCTTTTCTTTCAGCTAAGTAGTCTTACTGACCAATTTCTT 2414
Db 116 GTGATCTCGGAAGGATGCTTTTCTTTCAGCTAAGTAGTCTTACTGACCAATTTCTT 57
QY 2415 CTGGAACCAATAACATGCTCTTAAGCATTG 2444
Db 56 CTGGAACCAATAACATGCTCTTAAGCATTG 27
```

```
RESULT 10
BUI66874
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BUI66874. 867 bp mRNA linear EST 04-SEP-2002
AGENCOURT_7960371 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6166382
5', mRNA sequence.
BUI66874
BUI66874.1 GI:22680826
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 867)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTD/PTP
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL).
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM13526 row: j column: 15
High quality sequence stop: 579.
Location/Qualifiers
1..867
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6166382"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 72"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 2 kb. Library constructed by Life
Technologies."
```

FEATURES

source

```
RESULT 11
BUI82305
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
```

```
BUI82305
AGENCOURT_7840400 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:6146259
5', mRNA sequence.
BUI82305
BUI82305.1 GI:22696289
EST.
```

```
Query Match 24.1%; Score 785; DB 5; Length 867;
Best Local Similarity 96.1%; Pred. No. 6.1e-166;
Matches 827; Conservative 0; Mismatches 30; Indels 4; Gaps 2;

QY 1210 GAAGATGATGTAGATCAGGAACCTTGCAGACATACAGCCCTACGTGATAGATACCTTAAA 1269
Db 1 GAAGATGATGTAGATCAGGAACCTTGCAGACATACAGCCCTACGTGATAGATACCTTAAA 60
QY 1270 ACCAATGCCATATTGAGAATGGCCCATTTCTACTGGAGCAGCAACTGGAGATGAAGAT 1329
Db 61 ACCAATGCCATATTGAGAATGGCCCATTTCTCACTGGAGCAGCAACTGGAGATGAAGAT 120
QY 1330 GATGATGAAGAAGATCGCCCAAAATCTTGAGGAATATTAATCTTGATGAGCCAAATGCAGAA 1389
Db 121 GATGATGAAGAAGATCGCCCAAAATCTTGAGGAATATTAATCTTGATGAGCCAAATGCAGAA 180
QY 1390 AGTGATTTACATATAGCAGCTCCATGAAACAAATTCATGTTGAATGGCCAAATGGACGA 1449
Db 181 AGTGATTTACATATAGCAGCTCCATGAAACAAATTCATGTTGAATGGCCAAATGGACGA 240
QY 1450 CATAAATTTCCGAGTCCACAGTTCCACAGAGTTTCCACCTCGTTTCTCTGGATCCCTTA 1509
Db 241 CATAAATTTCCGAGTCCACAGTTCCACAGAGTTTCCACCTCGTTTCTCTGGATCCCTTA 300
QY 1510 GAACTGTGGCTCGGGCTCTGTGCTTCTGAGGATCACCCTTACTGACCAAGAGAG 1569
Db 301 GAACTGTGGCTCGGGCTCTGTGCTTCTGAGGATCACCCTTACTGACCAAGAGAG 360
QY 1570 AGCAGTCCATCCCATGACAGAAAGCAGACGGTTTTCAGCTCCAGTACTGGGGATTTGCCA 1629
Db 361 AGCAGTCCATCCCATGACAGAAAGCAGACGGTTTTCAGCTCCAGTACTGGGGATTTGCCA 420
QY 1630 AAAGCAAAACCCCGGCGAGCTGACTTGTGTGTAATCCCTGGATCCCGCGGAATGCAGAT 1689
Db 421 AAAGCAAAACCCCGGCGAGCTGACTTGTGTGTAATCCCTGGATCCCGCGGAATGCAGAT 480
QY 1690 AAAATTAGATGAAAAATGCTGACCTGGGAATGCTTGTGGGTGCATTAACACTTCCAG 1749
Db 481 AAAATTAGATGAAAAATGCTGACCTGGGAATGCTTGTGGGTGCATTAACACTTCCAG 540
QY 1750 GAAGACATCCAGACGCGTCAGTACCGCTCCATAGAGGTTTTAATAGGACGGGGTACAGC 1809
Db 541 GAAGACATCCAGACGCGTCAGTACCGCTCCATAGAGGTTTTAATAGGACGGGGTACAGC 600
QY 1810 ACCCTCGGGAATCTGGAGCAGCGCGTGTATGAGCTTTGAGCTGGCAACGGGAGATTAT 1869
Db 601 ACCCTCGGGAATCTGGAGCAGCGCGTGTATGAGCTTTGAGCTGGCAACGGGAGATTAT 660
QY 1870 TTGTTGAACCAATCTGGGGAAGACTATTCCAGAGACGAGACCATAGCCACATC 1929
Db 661 TTGTTGAACCAATCTGGGGAAGACTATTCCAGAGACGAGACCATAGCCACATC 720
QY 1930 ATAGAGCTGTAGGCAATTTCCAGGCACTTTTGTCTATCTGGAATAATTTCTCGGAA 1989
Db 721 ATAGAGCTGTAGGCAATTTCCAGGCACTTTTGTCTATCTTGAATAATTTCTCGGAA 780
QY 1990 TTCTTCAATCCGAGA--GGAGAACTCGCACATCAACAA--GCTGAAGCCCTTGAGCCT 2045
Db 781 TTCTTCCATCCAAAGGGAAAAATGCGACCATCATCCCAAGCTGGAAGCCCGGGAACCCC 840
QY 2046 CTTTGATGACTTGTGGAAA 2066
Db 841 CTTTGATGACTTGTGGGAAA 861
```

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 860)
NIH-MGC <http://mgc.ncl.nih.gov/>.
NATIONAL INSTITUTES OF HEALTH, MAMMALIAN GENE COLLECTION (MGC)
UNPUBLISHED (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM13474 row: d column: 04
High quality sequence stop: 676.
FEATURES
Location/Qualifiers
1..860
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6146259"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 67"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 23.5%; Score 765.2; DB 5; Length 860;
Best Local Similarity 97.9%; Pred. No. 1.8e-161;
Matches 797; Conservative 0; Mismatches 13; Indels 4; Gaps 2;
QY 1928 TCATGAGCTGCTAGGACGATATTCAGGACCTTGTCTATCTATCTGGAAATATTCGCG 1987
DB 1 TCATGAGCTGCTAGGACGATATTCAGGACCTTGTCTATCTATCTGGAAATATTCGCG 60
QY 1988 AATCTTCAATCGCAGAGAGAACTGCGACATCACCAGCTGAAGCCCTGAGCCTCT 2047
DB 61 AATCTTCAATCGCAGAGAGAACTGCGACATCACCAGCTGAAGCCCTGAGCCTCT 120
QY 2048 TTGATGTACTTGTGGAAGAGTATGGCTGCGCCCATGAAGATGCTGCACAGTTTACAGATT 2107
DB 121 TTGATGTACTTGTGGAAGAGTATGGCTGCGCCCATGAAGATGCTGCACAGTTTACAGATT 180
QY 2108 TCCTGATCCGATGTTAGAAATGGTTTCGAGAAACGAGCCTCAGCTGGCGAATGCTTTC 2167
DB 181 TCCTGATCCGATGTTAGAAATGGTTTCGAGAAACGAGCCTCAGCTGGCGAATGCTTTC 240
QY 2168 GGCATCTTGGTTGAATCTTAGCAAAATCTACCAATATTGCATTTCTGAGCTAGCAAAATG 2227
DB 241 GGCATCTTGGTTGAATCTTAGCAAAATCTACCAATATTGCATTTCTGAGCTAGCAAAATG 300
QY 2228 TTCCAGTACATTTGACACCTTAAACGGTGACTCTCATTTCTTTAACAGGATTTACAGTGAGCT 2287
DB 301 TTCCAGTACATTTGACACCTTAAACGGTGACTCTCATTTCTTTAACAGGATTTACAGTGAGCT 360
QY 2288 GGCCTCATCCTCAGACCTTTATTTGCTTTGAGGTAAGTCTGTTTGAACATTTTGTCTTTT 2347
DB 361 GGCCTCATCCTCAGACCTTTATTTGCTTTGAGGTAAGTCTGTTTGAACATTTTGTCTTTT 420
QY 2348 GTGCACTGTGATCCCTGGGAGAGGTAGTCTTTTGTCTTTCAGTAACTAGTCTTACTGACCA 2407
DB 421 GTGCACTGTGATCCCTGGGAGAGGTAGTCTTTTGTCTTTCAGTAACTAGTCTTACTGACCA 480
QY 2408 TTTTCTTCTCGGAAACAAATACATGCTCTAAGCATTTGTTTCTTGTGTTGTTGTTGTTGTTTCA 2467
DB 481 TTTTCTTCTCGGAAACAAATACATGCTCTAAGCATTTGTTTCTTGTGTTGTTGTTGTTGTTTCA 540

QY 2468 AATGTCATTTTTCGAATGAAAAATACCTTCCCTTGTGTTTGGCAGGTTTGTAACT 2527
DB 541 AATGTCATTTTTCGAATGAAAAATACCTTCCCTTGTGTTTGGCAGGTTTGTAACT 600
QY 2528 ATTTATGAAGAAATATTTTAGCTGAGTACTATATATATTTTACAATCTTAAGAAATTTATCAA 2587
DB 601 ATTTATGAAGAAATATTTTAGCTGAGTACTATATATATTTTACAATCTTAAGAAATTTATCAA 660
QY 2588 GTTGGGAACCAAGAAATAGCAAGGAATGTACAATTTTATCTTCTGCAAGGACAT 2647
DB 661 GTTGGGAACCAAGAAATAGCAAGGAATGTACAATTTTATCTTCTGCAAGGACAT 717
QY 2648 CATTCTCTATTTATAGTGTATGTAATGCACCTCTTAATGTTTACTTTCGATTAAATATG 2707
DB 718 CATTCTCTATTTATAGTGTATGTAATGCACCTCTTAATGTTTACTTTCATTAATATG 777
QY 2708 GGAGGGGGGACTCAAAATTTTCAGAAAGCTAAAAA 2741
DB 778 GGA-GGGGGACTCAAAATTTTCAGAAAGCTACCAA 810
RESULT 12
AUI24932
LOCUS AUI24932 NT2RM4 Homo sapiens cDNA clone NT2RM4000760 5', mRNA
DEFINITION AUI24932 828 bp mRNA linear EST 01-AUG-2002
sequence.
ACCESSION AUI24932
VERSION AUI24932.1 GI:10949648
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 828)
AUTHORS Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S., Saito, K., Yamamoto, J.,
Nakamura, Y., Nishikawa, T., Nagai, T., Suzuki, Y., Sugano, S. and
Isogai, T.
TITLE HRI human cDNA project (Ota, T., Wakamatsu, A., Ozawa, M., Ishii, S.,
Saito, K., Yamamoto, J., Nakamura, Y., Nishikawa, T., Nagai, T.,
Suzuki, Y., Sugano, S., Isogai, T.)
JOURNAL Unpublished (2000)
COMMENT Contact: Takeo Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
Location/Qualifiers
1..828
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NT2RM4000760"
/cell_type="teratocarcinoma"
/cell_line="NT2"
/clone_lib="NT2RM4"
/note="Vector: pME18SFL3; mRNA from uninduced NT2 neuronal
precursor cells"
ORIGIN
Query Match 23.3%; Score 759.2; DB 1; Length 828;
Best Local Similarity 98.8%; Pred. No. 4e-160;
Matches 795; Conservative 0; Mismatches 6; Indels 4; Gaps 3;
QY 1337 AAGAAGACTGCCCAATCTGAGGAAATATAATCTTGTAGGCAAAATGCAGAAAGTATT 1396
DB 14 AAGAAGACTGCCCAATCTGAGGAAATATAATCTTGTAGGCAAAATGCAGAAAGTATT 73

1397 ACACATATAGCAGCTCCTATGAACAATTCATGTTGAATTCGCAATTCGACGACATATAA 1456
|||||
74 ACACATATAGCAGCTCCTATGAACAATTCATGTTGAATTCGCAATTCGACGACATATAA 133
|||||
1457 TTCCCGAGTCACAGTTCACAGAGTTTCCACCTCGTTCTCTCTGGATTCCTAGAACCTG 1516
|||||
134 TTCCCGAGTCACAGTTCACAGAGTTTCCACCTCGTTCTCTCTGGATTCCTAGAACCTG 193
|||||
1517 TGGCTTGGGCTCTGTCTTCTGAGGATCACCACTTACTGAGCAAGAGAGCAGTC 1576
|||||
194 TGGCTTGGGCTCTGTCTTCTGAGGATCACCACTTACTGAGCAAGAGAGCAGTC 253
|||||
1577 CATCCCATGACAGAGCAGAGCGTTTCAGCTCCAGTCTGAGTTCGGGATTCGCCAAAGCAA 1636
|||||
254 CATCCCATGACAGAGCAGAGCGTTTCAGCTCCAGTCTGAGTTCGGGATTCGCCAAAGCAA 313
|||||
1637 AAACCCGGGCGAGCTGACTTGTGTGTAATCCCTCGATCCGCGGAATCGAGATAAATTA 1696
|||||
314 AAACCCGGGCGAGCTGACTTGTGTGTAATCCCTCGATCCGCGGAATCGAGATAAATTA 373
|||||
1697 GAGTAAATTCCTGACCTGGGAAATGCTTGTGGTGCATTAACACTTCACGGAAGACA 1756
|||||
374 GAGTAAATTCCTGACCTGGGAAATGCTTGTGGTGCATTAACACTTCACGGAAGACA 433
|||||
1757 TCCAGACGGCTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGACCCCTG 1816
|||||
434 TCCAGACGGCTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGACCCCTG 493
|||||
1817 CGGACATCTGGAGCAGCGCTGTATGGCAATTTGAGCTGGCAACGGAGATTAATTTGTTG 1876
|||||
494 CGGACATCTGGAGCAGCGCTGTATGGCAATTTGAGCTGGCAACGGAGATTAATTTGTTG 553
|||||
1877 ACCCATCTCTGGGAGAGACTATCCAGAGCAGAGACACATAGCCACATCATAGAGC 1936
|||||
554 ACCCATCTCTGGGAGAGACTATCCAGAGCAGAGACACATAGCCACATCATAGAGC 613
|||||
1937 TGTAGGCAATTCCTCAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGGAATTCCTCA 1996
|||||
614 TGTAGGCAATTCCTCAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGGAATTCCTCA 673
|||||
1997 ATCGCAGAGAGAACTGCGACACATCACAAGCTGAAGCCCTGGAGCCCTTTTGATGTAC 2056
|||||
674 ATCGCAGAGAGAACTGCGACACAT--ACCAACTGAAGCCCTGGA--CCTCTTTGATGTAC 730
|||||
2057 TTGTGAAAAGTATGCTGCCCCATGAAGATGCTGCACAGTTTACAGATTTCTGATCC 2116
|||||
731 TTGTGAAAAGTATGCTGCCCCATGAAGATGCTGCACAGTTTACAGATTTCTGATCC 789
|||||
2117 CGATGTTAGAAAATGGTTCCAGAAA 2141
|||||
790 CGATGTTAGAAAATGGNTCCAGAAA 814
|||||

RESULT 13

BU147088
LOCUS
DEFINITION
AGENCOURT 8727776 NIH_MGC_47 Homo sapiens cDNA clone IMAGE:6342174
5' mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BU147088
EST.
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 928)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLML)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLML at:

<http://image.llnl.gov>
Plate: LCM2541 row: g column: 07
High quality sequence stop: 619.

FEATURES
source

1..928
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6342174"
/tissue_type="neuroblastoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_47"
/note="Organ: brain; Vector: pOTB7; Site 1: XhoI; Site 2:
ECORI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adapter: GGCACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies).
Note: this is a NIH_MGC Library."

ORIGIN

Query Match 23.1%; Score 753; DB 5; Length 928;
Best Local Similarity 95.7%; Pred. No. 1e-158;
Matches 883; Conservative 0; Mismatches 25; Indels 15; Gaps 10;
QY 845 AAGCAGGTGCTCCTCCTTACAGGTCTGAGTACGGTCTCCACAGCAAACTTA 904
DB 1 AAGCAGGTGCTCCTCCTTACAGGTCTGAGTACGGTCTCCACAGCAAACTTA 60
QY 905 TAGAAAAATATCTAAAAACAAGAAAAAAGTAAAGAAAAAGAAAAAGAAAAAG 964
DB 61 TAGAAAAATATCTAAAAACAAGAAAAAAGTAAAGAAAAAGAAAAAGAAAAAG 120
QY 965 AGTTATTGGAGAACGCTGACGAGATAGAAAGAAATGGAGCGAGAACTGAAAGAAA 1024
DB 121 AGTTATTGGAGAACGCTGACGAGATAGAAAGAAATGGAGCGAGAACTGAAAGAAA 180
QY 1025 TAATAGAGAAAAATACACCTCAGCTGCACTTCCATGACGAGATGGCGAACTGCCC 1084
DB 181 TAATAGAGAAAAATACACCTCAGCTGCACTTCCATGACGAGATGGCGAACTGCCC 240
QY 1085 CAGAGTCAAACTAAAAACAACAGATTTAGAGGAGGGGCTGAGCGAGAGACTGCAAAAG 1144
DB 241 CAGAGTCAAACTAAAAACAACAGATTTAGAGGAGGGGCTGAGCGAGAGACTGCAAAAG 300
QY 1145 ACAATGGTGAAGCTGAGGACCAAGAGAAAGAGATGCTGAGAAAGAAACATTGAAA 1204
DB 301 ACAATGGTGAAGCTGAGGACCAAGAGAAAGAGATGCTGAGAAAGAAACATTGAAA 360
QY 1205 AAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
DB 361 AAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 1265 CTAAAAACCAATGGCCATATTTGAGATGGCCCACTTCTCACTGGAGCAGCAACTGGA 1324
DB 421 CTAAAAACCAATGGCCATATTTGAGATGGCCCACTTCTCACTGGAGCAGCAACTGGA 480
QY 1325 AAGATGATGATGAAGAGACTGCCCAATCTCTGAGAAATATAATCTTGTATGAGCCAAATG 1384
DB 481 AAGATGATGATGAAGAGACTGCCCAATCTCTGAGAAATATAATCTTGTATGAGCCAAATG 540
QY 1385 CAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1444
DB 541 CAGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 1445 GACGACATAAAATTTCCCGAGTCACAGTT--CCGAGAGTTTTCACCTCGTTGTTCTCTGGA 1503

601 GAGGACATAAATCCCGAGTACAGTCCCGAGAGTTTCCACCTCGTGTCTCTGGA 660
1504 TCCTTAGAACCTGTGGCCCTGCGGCTCTGTGCTTTCTGAGGGATCACCACCTTACTGAGCAA 1563
661 TCCTTAGAACCTGTGGCCCTGCGGCTCTGTGCTTTCTGAGGGATCACCACCTTACTGAGCAA 720
1564 GA-GGAGAGCAGTCCATCCATGACAGAGACGAGAC-GGTTTCAGCCTCCAG--TACTGG 1619
721 GAGGAGAGCAGTCCATCCCTGACAGAGCCGAGCGGTTTTCAGCCTCCAGGTACTGGG 780
1620 GGATTTCGCAAAA--GCAAAAACCGGCGAGCTGACTTCTT--GGTGAATCCCTGGATC 1675
781 GGATTTCGCAAAAACGAAACCCCGGCGAGCTGACCGGTTTGTGTGATCCCTTGGATC 840
1676 CCG-GGAATGAGATAAAA--TTAGATTAATAATG-CTGACCTGGGAAA---TGCTTGT 1729
841 CCGGGAAGCAGATAAATTTAGAGTAATAAATTCCTGACCTGAGAAAATGCCTTGGTG 900
1730 GGGTGCATAAACACTTCAGGAA 1752
901 GGGGGCATAAACCTTTCGGAA 923

RESULT 14
BX423898 900 bp mRNA linear EST 01-MAY-2004
LOCUS
DEFINITION BX423898 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
CSODN003YA14 5-PRIME, mRNA sequence.

ACCESSION BX423898
VERSION
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 900)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 13, 2003 this sequence version replaced gi:30643492.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 4027.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CSIAN0012E08QP1&c=4027.r.

Location/Qualifiers
1..900
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODN003YA14"
/tissue_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Query Match 22.8%; Score 741.8; DB 5; Length 900;
Best Local Similarity 97.0%; Pred. No. 3,5e-156;
Matches 808; Conservative 2; Mismatches 14; Indels 9; Gaps 5;

QY 910 AATAATCTAAAAACAAAAAAGAAAACTGAAAAAGAAACAGAGAGCGAGCTGAGTTA 969
DB 5 AATAATCTAAAAACAAAAAAGAAAACTGAAAAAGAAACAGAGAGCGAGCTGAGTTA 64
QY 970 TTGGAGAGC--GCCTGCAGGAGATAGAGAAATTGGAGCGAGAAAGCTGAAAGGAAATAA 1027
DB 65 TTGGAGAGCAGCGCCCTGCAGGAGATAGAGAAATTGGAGCGAGAAAGCTGAAAGGAAATAA 124
QY 1028 TAGAAGAAAAATCACTCTGAGCTGCACTTCCATGACAGAGATGCGGATGCTGCCAG 1087
DB 125 TAGAAGAAAAATCACTCTGAGCTGCACTTCCATGACAGAGATGCGGATGCTGCCAG 184
QY 1088 AGGTCAACTTAAAAACAAACAGGATTTAGAGAGCGGCTGAGGACAGACTGCAAAAGACA 1147
DB 185 AGGTCAACTTAAAAACAAACAGGATTTAGAGAGCGGCTGAGGACAGACTGCAAAAGACA 244
QY 1148 ATGTGTGAAGCTGAGGACCGAGAGAGAAAGATGCTGAGAGAAAGAAACATTTGAAAAAG 1207
DB 245 ATGTGTGAAGCTGAGGACCGAGAGAGAAAGATGCTGAGAGAAAGAAACATTTGAAAAAG 304
QY 1208 ATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1267
DB 305 ATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 364
QY 1268 AACCACATGCGCATATTTGAGATGCGCCATTTCTCACTGAGGACGCAACTGGACGATGAAG 1327
DB 365 AACCACATGCGCATATTTGAGATGCGCCATTTCTCACTGAGGACGCAACTGGACGATGAAG 424
QY 1328 ATGATGATGAAG 1387
DB 425 ATGATGATGAAG 484
QY 1388 AAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
DB 485 AAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 544
QY 1447 CGACATATAATTTCCGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 1506
DB 545 CGACATATAATTTCCGAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCC 604
QY 1507 TTAGAACCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 1566
DB 605 TTAGAACCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCT 664
QY 1567 GAGACAGTCCATCCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1626
DB 665 GAGACAGTCCATCCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 724
QY 1627 CCAAGACAAAAACCGGCGAGCTGACCTGTTGTGTAATCCCTGGATCCGCGGATGCA 1686
DB 725 CCAAGACAAAAACCGGCGAGCTGACCTGTTGTGTAATCCCTGGATCCGCGGATGCA 782
QY 1687 GATAAATTTAGAGTAAAAAATTCCTGACCTGGGAAATGCTTGTGGTGCATAA 1739
DB 783 GATAAATTTAGAGTAAAAAATTTG---ACTKGGGAATGCTTGTGGTGCATAA 831

RESULT 15
BX442846

LOCUS
DEFINITION BX442846 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSODF038Y110 5-PRIME, mRNA sequence.

ACCESSION BX442846
VERSION BX442846.2 GI:47037734

KEYWORDS
SOURCE EST.

ORGANISM Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 960)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)

On May 22, 2003 this sequence version replaced gi:31018541.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dr) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 This sequence belongs to sequence cluster 4027.r
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?e=CS0DF038BE05QPI&c=4027.r.
 Location/Qualifiers
 1..960
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DF038Y110"
 /tissue_type="FETAL BRAIN"
 /dev_stage="fetal"
 /clone_lib="Homo sapiens FETAL BRAIN"
 /note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo (dr) primer. Five prime end
 enriched, double-strand cDNA was digested with Not I and
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6
 vector. Library was not normalized."

FEATURES
 source

ORIGIN

Query Match 22.7%; Score 739; DB 5; Length 960;
 Best Local Similarity 94.5%; Pred. No. 1.5e-155;
 Matches 757; Conservative 25; Mismatches 16; Indels 3; Gaps 2;

QY 153 TCAGAAAGCCGGAGCCTCAACAGAAAGCTCTTCTAGTTCTCTCTCCACCGCCACC 212
 DB 154 TTCAAGAAAGCCGGAGCCTCAACAGAAAGCTCTTCTAGTTCTCTCTCCACCGCCACC 213

QY 213 ACCACACCC--ACCGCCACCTTTGCCAGACCCACACCCCGGAGCAGAGGAGATC 270
 DB 214 ACCACACCCAGCCGCCACCTTTGCCAGACCCACACCCCGGAGCAGAGGAGATC 273

QY 271 CTGGGATCAGATGATGAGGAGCAAGAGGACCCCTGCGGACTACTGCAAGGTGGATATCAT 330
 DB 274 CTGGGATCAGATGATGAGGAGCAAGAGGACCCCTGCGGACTACTGCAAGGTGGATATCAT 333

QY 331 CCAAGTGAATTCGAGACCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGG 390
 DB 334 CCAAGTGAATTCGAGACCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGG 393

QY 391 GGGCACCTTCTACTGTCTGCTGTGCTGGGATATGCGGGGAAAGATTTGTTGCAATG 450
 DB 394 GGGCACCTTCTACTGTCTGCTGTGCTGGGATATGCGGGGAAAGATTTGTTGCAATG 453

QY 451 AAAGTTGTAAGAGTGGCCAGCATTTATACGAGACAGCCTTGGATGAATTAATTTGCTC 510
 DB 454 AAAGTTGTAAGAGTGGCCAGCATTTATACGAGACAGCCTTGGATGAATTAATTTGCTC 513

QY 511 AAATGCTTCGAGAAAGTATCCAGTGACCCAAACAAAGACATGTGTGGTCCAGCTCAT 570
 DB 514 AAATGCTTCGAGAAAGTATCCAGTGACCCAAACAAAGACATGTGTGGTCCAGCTCAT 573

QY 571 GACGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCAATGGTCTTCGAAGTACT 630
 DB 574 GACGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCAATGGTCTTCGAAGTACT 633

QY 631 GGCCACATCTCTCAAGTGGATCAATCAATCCAACTATCAAGGCCTCCCAAGTACGTTGT 690
 DB 634 GGCCACATCTCTCAAGTGGATCAATCAATCCAACTATCAAGGCCTCCCAAGTACGTTGT 693

QY 691 GTGAAGAGTATCATTTCCAGAGGTCCTTCAAGGGTTAGATTACTTTACACAGTAAGTCAAG 750
 DB 694 GTGAAGAGTATCATTTCCAGAGGTCCTTCAAGGGTTAGATTACTTTACACAGTAAGTCAAG 753

751 ATCATTTCATCTGACATATAAAGCGGAAAAATATCTTGAATCTGTGTGATGATCATATCTG 810
 754 ATCATTTCATCTGACATATAAAGCGGAAAAATATCTTGAATCTGTGTGATGATCATATCTG 813
 811 AGAAGATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTTCA-GG 869
 814 AGAAGATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTTCAARRG 873
 870 GTCTGCAGTGTGATGACGGCTCCACAGCAGAAACCTATAGGAAAAAATATCTTAAAAACAAAA 929
 874 GTMTCAGTGTGATGACGGCTCCACAGCAGAAACCTATAGGAAAAAATATCTTAAAAACAAAA 933
 930 GAAAAAACTGAAAAAGAACACA 950
 934 RAAAAACYKAAAAAGAACARA 954

Search completed: January 23, 2005, 11:04:28
 Job time : 9769.26 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: January 23, 2005, 11:04:34 ; Search time 1666.02 Seconds
(without alignments)
11219.162 Million cell updates/sec

Title: US-10-799-676-1
Perfect score: 3253
Sequence: 1 tcggcgagcgagtgaggc.....ctatataattactattaaa 3253

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues
Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3253	100.0	3253	9	US-09-759-359A-1
2	3253	100.0	3253	15	US-10-207-973-1
3	3253	100.0	3253	17	US-10-799-676-1
4	2533.8	77.9	3715	18	US-10-618-941-36
5	2533.8	77.9	4698	14	US-10-198-846-10286
6	2529	77.7	3699	16	US-10-425-114-16424
7	2497	76.8	3745	10	US-09-981-397A-21
8	666.2	20.5	90541	9	US-09-759-359A-3
9	666.2	20.5	90541	15	US-10-207-973-3
10	666.2	20.5	90541	17	US-10-799-676-3
11	632.8	19.5	4326	9	US-09-880-107-3294
12	632.8	19.5	4326	15	US-10-172-118-875

13	632.8	19.5	4326	16	US-10-342-887-875	Sequence 875, App
14	629.6	19.4	4349	15	US-10-252-157-374	Sequence 374, App
15	629.6	19.4	4639	10	US-09-814-353-21987	Sequence 21987, A
16	629.4	19.3	923	16	US-10-305-720-703	Sequence 703, App
17	488.6	15.0	574	15	US-10-029-386-9677	Sequence 9677, App
c 18	466	14.3	466	9	US-09-864-761-22399	Sequence 22399, A
c 19	441.8	13.6	533	15	US-10-029-386-1525	Sequence 1525, Ap
20	441.8	13.6	533	15	US-10-029-386-12953	Sequence 12953, A
21	404.2	12.4	1835	15	US-10-353-690-47	Sequence 47, Appl
22	404.2	12.4	1835	18	US-10-723-860-1548	Sequence 1548, Ap
23	404.2	12.4	1911	18	US-10-723-860-5975	Sequence 5975, Ap
24	399.8	12.3	421	10	US-09-918-995-33414	Sequence 33414, A
25	362.4	11.1	1112	10	US-09-814-353-20219	Sequence 20219, A
26	359.6	11.1	611	18	US-10-425-115-78853	Sequence 78853, A
c 27	357.8	11.0	466	9	US-09-864-761-5631	Sequence 5631, Ap
28	347.8	10.7	425	9	US-09-960-352-2889	Sequence 2889, Ap
29	338.6	10.4	716	9	US-09-728-445-641	Sequence 641, App
30	293.4	9.0	1427	9	US-09-870-962-10	Sequence 10, Appl
c 31	293	9.0	617	16	US-10-152-319A-216	Sequence 216, App
32	272	8.4	332	14	US-10-198-846-10124	Sequence 10124, A
33	271	8.3	387	14	US-10-198-846-8302	Sequence 8302, Ap
34	254	7.8	793	16	US-10-305-720-79	Sequence 79, Appl
35	248	7.6	249	16	US-10-305-720-808	Sequence 808, App
c 36	193.6	6.0	571	14	US-10-198-846-134	Sequence 134, App
37	184.4	5.7	1933	16	US-10-425-114-35327	Sequence 35327, A
38	184.4	5.7	2267	18	US-10-425-115-152692	Sequence 152692, A
39	181.4	5.6	2128	16	US-10-425-114-8287	Sequence 8287, Ap
40	180	5.5	346	10	US-09-814-353-1961	Sequence 1961, Ap
41	180	5.5	346	10	US-09-814-353-8308	Sequence 8308, Ap
42	179	5.5	340	10	US-09-814-353-14692	Sequence 14692, A
43	177.2	5.4	617	10	US-09-871-161-396	Sequence 396, App
44	177.2	5.4	2842	16	US-10-424-599-72889	Sequence 72889, A
45	175.2	5.4	1076	16	US-10-424-599-39344	Sequence 39344, A

ALIGNMENTS

RESULT 1
US-09-759-359A-1
; Sequence 1, Application US/09759359A
; Patent No. US20020094560A1
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/09/759,359A
; CURRENT FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3253
; TYPE: DNA
; ORGANISM: Human
US-09-759-359A-1

Query Match	100.0%	Score 3253;	DB 9;	Length 3253;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3253;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	TCGGCGGAGCGAGTGGAGCTCAGCCAGCTCGTTCGGCGCCCGCGTGGCGTTCGCGA	60	
Db	1	TCGGCGGAGCGAGTGGAGCTCAGCCAGCTCGTTCGGCGCCCGCGTGGCGTTCGCGA	60	
QY	61	AGCCCCCGCCCGCTTCGGCGCTCGGAATGAGCTCCGGAAGTGTGGCCATTTCAG	120	
Db	61	AGCCCCCGCCCGCTTCGGCGCTCGGAATGAGCTCCGGAAGTGTGGCCATTTCAG	120	
QY	121	GCCCGAAGCGGAGCCGGAAGAGAGAACTCCGAAAAGCCGAGGCTCAACAGAAA	180	
Db	121	GCCCGAAGCGGAGCCGGAAGAGAGAACTCCGAAAAGCCGAGGCTCAACAGAAA	180	

2341 GCCTTTTGTGCACTGTGATCCCTGGGAAGGTAGTCTTTTGTCTTCAGCTAGTAGTTTA 2400
2401 CTGACCAATTTCTTCTGGAAACAATAACATGCTCTTAAGCAATTTCTTGTGTTGTG 2460
2401 CTGACCAATTTCTTCTGGAAACAATAACATGCTCTTAAGCAATTTCTTGTGTTGTG 2460
2461 ACATTCAAAATGTCATTTTCTTGAATGAAATAATCTTCCCTTTGTGTTTGGCAGGTTT 2520
2461 ACATTCAAAATGTCATTTTCTTGAATGAAATAATCTTCCCTTTGTGTTTGGCAGGTTT 2520
2521 TGTAACTATTATTGAAGAAATATTTTAGCTGAGTACTATATATTTTACAACTCTTAAGAAA 2580
2521 TGTAACTATTATTGAAGAAATATTTTAGCTGAGTACTATATATTTTACAACTCTTAAGAAA 2580
2581 TTATCAAGTTGGAAACCAAGAAATAGCAAGGAAATGTACAAATTTTATCTTCTGGCAAA 2640
2581 TTATCAAGTTGGAAACCAAGAAATAGCAAGGAAATGTACAAATTTTATCTTCTGGCAAA 2640
2641 GGGACATCATCTCTGTATTATAGTGTATGTAATGCAACCTGTAAATGTTTACCTTTGGATT 2700
2641 GGGACATCATCTCTGTATTATAGTGTATGTAATGCAACCTGTAAATGTTTACCTTTGGATT 2700
2701 AAATATGCGAGGGGAGCTCAAAATTTTCAAGAAAGCTTAAAGGAAATGTAATAAGG 2760
2701 AAATATGCGAGGGGAGCTCAAAATTTTCAAGAAAGCTTAAAGGAAATGTAATAAGG 2760
2761 AAAAATPACTCTTATATTAATAACCTTTCTTTGTTTTTTTGTGTTTTTCTTATTTTCAATTT 2820
2761 AAAAATPACTCTTATATTAATAACCTTTCTTTGTTTTTTTGTGTTTTTCTTATTTTCAATTT 2820
2821 ATTAAATACACTTAACTGTCGAAAGCACTATGAAAGGAAATTAATACCATGAAAGGATCA 2880
2821 ATTAAATACACTTAACTGTCGAAAGCACTATGAAAGGAAATTAATACCATGAAAGGATCA 2880
2881 AAAATCATAAATCAAAACCCCTATAGTCCAAAGCAATTCATCTCGCGGCTCAACTT 2940
2881 AAAATCATAAATCAAAACCCCTATAGTCCAAAGCAATTCATCTCGCGGCTCAACTT 2940
2941 TTTAAATCTTATPACTAGTACCTGAGACTCTGCTGCTCAATATTAATTTCTAAATCTTAC 3000
2941 TTTAAATCTTATPACTAGTACCTGAGACTCTGCTGCTCAATATTAATTTCTAAATCTTAC 3000
3001 CACCAAGTTAGGCCCGTAATGCTCTCTCTGTAATCTGTGTAATCTGTGTAATCTGTGTTTCT 3060
3001 CACCAAGTTAGGCCCGTAATGCTCTCTCTGTAATCTGTGTAATCTGTGTAATCTGTGTTTCT 3060
3061 ATTTATTTAGTGGGCTCTGTTTATCTTTCGCCACATCTTGTTCACATTTTCTAGTTA 3120
3061 ATTTATTTAGTGGGCTCTGTTTATCTTTCGCCACATCTTGTTCACATTTTCTAGTTA 3120
3121 CTCTTATCTTTGGGCTGATTAATCTTCTCATTTATCTCATATAAATCTTCTGAATTTTTC 3180
3121 CTCTTATCTTTGGGCTGATTAATCTTCTCATTTATCTCATATAAATCTTCTGAATTTTTC 3180
3181 ACATAAACTACTAGAGCTACCTCAGCAATCTCTGTTTTTAAACGCGAGGATTAATATA 3240
3181 ACATAAACTACTAGAGCTACCTCACCATCTCTGTTTTTAAACGCGAGGATTAATATA 3240
3241 ATTACTATTAAA 3253
3241 ATTACTATTAAA 3253

RESULT 2
US-10-207-973-1
; Sequence 1, Application US/10207973
; Publication No. US20030175927A1
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF

FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/10/207,973
; CURRENT FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3253
; TYPE: DNA
; ORGANISM: Human
; US-10-207-973-1

Query Match 100.0%; Score 3253; DB 15; Length 3253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGCGGAGCGAGTGGAGCTGCAGCCAGCTCGTCTCGGCGCCCGCGTCCGCGTCCGCG 60
DB 1 TCGCGGAGCGAGTGGAGCTGCAGCCAGCTCGTCTCGGCGCCCGCGTCCGCGTCCGCG 60

QY 61 AGCCCCCGCCCGCTTCGCGCGCTCGGAATGAGCTCCGGAAGTCTGCGCATTCAG 120
DB 61 AGCCCCCGCCCGCTTCGCGCGCTCGGAATGAGCTCCGGAAGTCTGCGCATTCAG 120

QY 121 GCCCGAAAGCGAGGCGGAAAGAGAGAAACATCCGAAAGCGGAGGCTCAACAGAAA 180
DB 121 GCCCGAAAGCGAGGCGGAAAGAGAGAAACATCCGAAAGCGGAGGCTCAACAGAAA 180

QY 181 GCTCCTTAGTTTCTCTCTCCACGCCCAACCAACCCAGCCACCTTTTGCAGAC 240
DB 181 GCTCCTTAGTTTCTCTCTCCACGCCCAACCAACCCAGCCACCTTTTGCAGAC 240

QY 241 CCCACACCCCGGAGGAGGAGGAGATCTTGGGATCAGATGATGAGGAGCAAGGAC 300
DB 241 CCCACACCCCGGAGGAGGAGGAGATCTTGGGATCAGATGATGAGGAGCAAGGAC 300

QY 301 CTTGGGAGCTACTGCAAAAGTGGATATCATCCAGTCAAAATTTGGAGACCTCTTCAATGGC 360
DB 301 CTTGGGAGCTACTGCAAAAGTGGATATCATCCAGTCAAAATTTGGAGACCTCTTCAATGGC 360

QY 361 CGGTATCATGTTATTAGAAAGCTTGGATGGGGGCACTCTCTACTGTCTGGCTGTCTGG 420
DB 361 CGGTATCATGTTATTAGAAAGCTTGGATGGGGGCACTCTCTACTGTCTGGCTGTCTGG 420

QY 421 GATATGAGGGGAAAGATTTGTTGCAATGAAAGTTGTAAGTCCAGCATTAATACG 480
DB 421 GATATGAGGGGAAAGATTTGTTGCAATGAAAGTTGTAAGTCCAGCATTAATACG 480

QY 481 GAGACAGCCTTCGATGAAATAAATTTGCTCAAAATGTTTCGAGAAAGTATCCAGTGAC 540
DB 481 GAGACAGCCTTCGATGAAATAAATTTGCTCAAAATGTTTCGAGAAAGTATCCAGTGAC 540

QY 541 CCAACAAAGACATGGTGTCCAGCTCATTGACGACTTCAAGATTTTCAAGATTTTCAAGTGGG 600
DB 541 CCAACAAAGACATGGTGTCCAGCTCATTGACGACTTCAAGATTTTCAAGATTTTCAAGTGGG 600

QY 601 ATACATGTCGATGTTGTTGAAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAA 660
DB 601 ATACATGTCGATGTTGTTGAAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAA 660

QY 661 TCCAACTATCAAGGCTCCCGAGTACGTTGTTGTAAGATATCATTCGAGAGTCTTCAA 720
DB 661 TCCAACTATCAAGGCTCCCGAGTACGTTGTTGTAAGATATCATTCGAGAGTCTTCAA 720

QY 721 GGGTTAGATTACTTACACAGTAAGTCAAGATCAATTCATCTGACATAAGCCGGAAT 780
DB 721 GGGTTAGATTACTTACACAGTAAGTCAAGATCAATTCATCTGACATAAGCCGGAAT 780

QY 781 ATCTTGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 ATCTTGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840

QY 841 CAGAAAGCAGGTTCT 900
DB 841 CAGAAAGCAGGTTCT 900

Db 841 CAGAAGCAGGTGCTCCTCTCTCAGGCTCGAGTGAGTACGGCTCCACAGAGAA 900
 Qy 901 CTTATAGGAAAAATATCTAAACAAAGAAAACTGAAAAAGAAACAGAGAGGCGAG 960
 Db 901 CTTATAGGAAAAATATCTAAACAAAGAAAACTGAAAAAGAAACAGAGAGGCGAG 960
 Qy 961 GCTGAGTATTGAGAGCGCTCGAGAGGATAGAGAAATGGAGCGAGAGAGCTGAAAGG 1020
 Db 961 GCTGAGTATTGAGAGCGCTCGAGAGGATAGAGAAATGGAGCGAGAGAGCTGAAAGG 1020
 Qy 1021 AAAATAATAGAGAAAAACATCACTCAGCTGCACTCTCAATGACCAAGATGGCGAATAC 1080
 Db 1021 AAAATAATAGAGAAAAACATCACTCAGCTGCACTCTCAATGACCAAGATGGCGAATAC 1080
 Qy 1081 TGCCAGAGGTGAATCTAAAAACAACAGATTTAGAGGAGGCGCTGAGCGAGAGAGCTGCA 1140
 Db 1081 TGCCAGAGGTGAATCTAAAAACAACAGATTTAGAGGAGGCGCTGAGCGAGAGAGCTGCA 1140
 Qy 1141 AAGGCAATGGTGAAGCTCAGGACCAAGAGAGAAAGAGATGCTGAGAAAGAAAAACATT 1200
 Db 1141 AAGGCAATGGTGAAGCTCAGGACCAAGAGAGAAAGAGATGCTGAGAAAGAAAAACATT 1200
 Qy 1201 GAAAAAGATGAAGATGATGATAGATCAGGAACTTGGCGAATAGACCTACGTGGATAGAA 1260
 Db 1201 GAAAAAGATGAAGATGATGATAGATCAGGAACTTGGCGAATAGACCTACGTGGATAGAA 1260
 Qy 1261 TCACCTAAACCAATGGCCATATTGAGATGGCCATCTCACTGGAGCAGCAACTGGAC 1320
 Db 1261 TCACCTAAACCAATGGCCATATTGAGATGGCCATCTCACTGGAGCAGCAACTGGAC 1320
 Qy 1321 GATGAAGATGATGATGAAGAGACTGCCCAATCTCTGAGGAATATAATCTTGATAGGCA 1380
 Db 1321 GATGAAGATGATGATGAAGAGACTGCCCAATCTCTGAGGAATATAATCTTGATAGGCA 1380
 Qy 1381 AATGCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 Db 1381 AATGCGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 Qy 1441 AATGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 Db 1441 AATGGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 Qy 1501 GATGCTTTAGAACCTGTGGCTCGGCTGTGCTTTCTGAGGATGATGATGATGATGATGAT 1560
 Db 1501 GATGCTTTAGAACCTGTGGCTCGGCTGTGCTTTCTGAGGATGATGATGATGATGATGAT 1560
 Qy 1561 CAAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Db 1561 CAAGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Qy 1621 GATTTGCAAAAGCAAAACCCGGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1680
 Db 1621 GATTTGCAAAAGCAAAACCCGGGAGCTGATGATGATGATGATGATGATGATGATGATGAT 1680
 Qy 1681 AATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Db 1681 AATCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Qy 1741 CACTTCCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 Db 1741 CACTTCCAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 Qy 1801 GGGTACAGACCCCTGGGAGATCTGGAGCAAGCGCTGATGATGATGATGATGATGATGATGAT 1860
 Db 1801 GGGTACAGACCCCTGGGAGATCTGGAGCAAGCGCTGATGATGATGATGATGATGATGATGAT 1860
 Qy 1861 GGAGATTTATTTGTTGAACCACTTCTGGGAGAGATTTCCAGAGAGAGAGAGAGAGAGAG 1920
 Db 1861 GGAGATTTATTTGTTGAACCACTTCTGGGAGAGATTTCCAGAGAGAGAGAGAGAGAGAG 1920
 Qy 1921 GCGGACATCATAGAGCTCTAGGAGATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1980
 Db 1921 GCGGACATCATAGAGCTCTAGGAGATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1980

Qy 1981 TCTCGGAAATCTTCAATCGCAGAGAGAACTGCGACA CATCA CCAAGCTGAAGCCCTGG 2040
 Db 1981 TCTCGGAAATCTTCAATCGCAGAGAGAACTGCGACA CATCA CCAAGCTGAAGCCCTGG 2040
 Qy 2041 AGCCTCTTGTGATGATCTTGTGAAAAAGTATGCTGCGCCCATGAAGATGCTGCACAGTTT 2100
 Db 2041 AGCCTCTTGTGATGATCTTGTGAAAAAGTATGCTGCGCCCATGAAGATGCTGCACAGTTT 2100
 Qy 2101 ACAGATTTCTGATCCGATCCGATGTTAGAAATGTTTCAGAAAAACGAGCCTCAGCTGGCGAA 2160
 Db 2101 ACAGATTTCTGATCCGATGTTAGAAATGTTTCAGAAAAACGAGCCTCAGCTGGCGAA 2160
 Qy 2161 TGCCTTCGGCATCTTGTGTTGAAATCTTAGCAAAATCTACCAATATTTGCAATCTTGAGCTA 2220
 Db 2161 TGCCTTCGGCATCTTGTGTTGAAATCTTAGCAAAATCTACCAATATTTGCAATCTTGAGCTA 2220
 Qy 2221 GCAATGTTCCAGTACATTTGGACCTAAACGGTGACTCTCATCTTTTAAACAGGATTAACA 2280
 Db 2221 GCAATGTTCCAGTACATTTGGACCTAAACGGTGACTCTCATCTTTTAAACAGGATTAACA 2280
 Qy 2281 GTGAGCTGGCTTCACTCCTCAGACCTTTATTTTCTTTGAGGTACTGTTGTTGACATTTT 2340
 Db 2281 GTGAGCTGGCTTCACTCCTCAGACCTTTATTTTCTTTGAGGTACTGTTGTTGACATTTT 2340
 Qy 2341 GCTTTTGTGCACTGTGATCCTCGGGAAGGTAGTCTTTTGTCTTCAAGCTAAGTAGTTTA 2400
 Db 2341 GCTTTTGTGCACTGTGATCCTCGGGAAGGTAGTCTTTTGTCTTCAAGCTAAGTAGTTTA 2400
 Qy 2401 CTGACCAATTTCTTCTCGAAACAAATACATGCTCTAAAGCAATGTTTCTTTGTTGTGTG 2460
 Db 2401 CTGACCAATTTCTTCTCGAAACAAATACATGCTCTAAAGCAATGTTTCTTTGTTGTGTG 2460
 Qy 2461 ACATTTCAAAATGCTATTTTGTGAAATGAAATACTTTCCCTTTGTTGTTGTTGTTGTT 2520
 Db 2461 ACATTTCAAAATGCTATTTTGTGAAATGAAATACTTTCCCTTTGTTGTTGTTGTTGTT 2520
 Qy 2521 TGTAACTATTTATGAAGAAATTTTGTAGCTGAGTACTATATAATTAACAATCTTTAAGAA 2580
 Db 2521 TGTAACTATTTATGAAGAAATTTTGTAGCTGAGTACTATATAATTAACAATCTTTAAGAA 2580
 Qy 2581 TTATCAAGTTCGGGAACCAAGAAAAATAGCAAGGAAATGTACAATTTTATCTTTGCGAAA 2640
 Db 2581 TTATCAAGTTCGGGAACCAAGAAAAATAGCAAGGAAATGTACAATTTTATCTTTGCGAAA 2640
 Qy 2641 GGGACATCAATCTCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 Db 2641 GGGACATCAATCTCTGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2700
 Qy 2701 AAATATGGAGGGGAGCTCAAAATTCAGAAAAAGCTAAAAAAGTAAATTAAGTAAAG 2760
 Db 2701 AAATATGGAGGGGAGCTCAAAATTCAGAAAAAGCTAAAAAAGTAAATTAAGTAAAG 2760
 Qy 2761 AAAAATACTCTTATATTAATAAATACCCCTTTCTTTGTTTGTGTTTGTGTTTGTGTTT 2820
 Db 2761 AAAAATACTCTTATATTAATAAATACCCCTTTCTTTGTTTGTGTTTGTGTTTGTGTTT 2820
 Qy 2821 ATTAATACTCTTATATTAATAAATACCCCTTTCTTTGTTTGTGTTTGTGTTTGTGTTT 2880
 Db 2821 ATTAATACTCTTATATTAATAAATACCCCTTTCTTTGTTTGTGTTTGTGTTTGTGTTT 2880
 Qy 2881 AAAATCATAAATCAAAAACCCCACTATAGTCCAAACGACAAATTCATTCTCGGGGTCAACT 2940
 Db 2881 AAAATCATAAATCAAAAACCCCACTATAGTCCAAACGACAAATTCATTCTCGGGGTCAACT 2940
 Qy 2941 TTTAAACATCTTATATTAATAAATACCCCTTTCTTTGTTTGTGTTTGTGTTTGTGTTT 3000
 Db 2941 TTTAAACATCTTATATTAATAAATACCCCTTTCTTTGTTTGTGTTTGTGTTTGTGTTT 3000
 Qy 3001 CACCAAGTTAGGCGGCTGCT 3060
 Db 3001 CACCAAGTTAGGCGGCTGCT 3060

QY 3061 ATTATTAGTGGCTCGTTTATCTTTTCGCCACATCTTTGTTCACTATTTTCTAGTTA 3120
DB 3061 ATTATTAGTGGCTCGTTTATCTTTTCGCCACATCTTTGTTCACTATTTTCTAGTTA 3120
QY 3121 CTCCTATCTTTGGCTGATTAATCTCTCTCATPATCTCATATAAACTTCTGAATTTTTC 3180
DB 3121 CTCCTATCTTTGGCTGATTAATCTCTCTCATPATCTCATATAAACTTCTGAATTTTTC 3180
QY 3181 ACATAAACTACTAGAGCTACCTCACCATCTCTGTTTTTAAACCGAGGAGTTACTATATA 3240
DB 3181 ACATAAACTACTAGAGCTACCTCACCATCTCTGTTTTTAAACCGAGGAGTTACTATATA 3240
QY 3241 ATTACTATTAAA 3253
DB 3241 ATTACTATTAAA 3253

RESULT 3

US-10-799-676-1
; Sequence 1, Application US/10799676
; Publication No. US20040157297A1
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001043DIV II
; CURRENT APPLICATION NUMBER: US/10/799,676
; CURRENT FILING DATE: 2004-03-15
; PRIOR APPLICATION NUMBER: 10/207,973
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/759,359
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3253
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-799-676-1

Query Match 100.0%; Score 3253; DB 17; Length 3253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCGCGGAGCGAGTGAGGCTGCAGCCAGCTCGTCTCGGCGCCGCGTCCGCGTCCGCA 60
DB 1 TCGCGGAGCGAGTGAGGCTGCAGCCAGCTCGTCTCGGCGCCGCGTCCGCGTCCGCA 60
QY 61 AGCCCCCGCCCGCTTCGCGCGCTCGGAATGAGCTCCGGAAGTGTGGCCATTTCAG 120
DB 61 AGCCCCCGCCCGCTTCGCGCGCTCGGAATGAGCTCCGGAAGTGTGGCCATTTCAG 120
QY 121 GCCCGAAGCGGAGCGGCGGAAAGAGAGAAACATCCGAAAGCCGAGCCTCAACAGAAA 180
DB 121 GCCCGAAGCGGAGCGGCGGAAAGAGAGAAACATCCGAAAGCCGAGCCTCAACAGAAA 180
QY 181 GCTCTTTTATGTTCTCTCTCCACCACCAACACCAACCGCCACTTTTGGCAGAC 240
DB 181 GCTCTTTTATGTTCTCTCTCCACCACCAACACCAACCGCCACTTTTGGCAGAC 240
QY 241 CCCACACCCCGGAGCCAGAGGAGATCTTGGGATCAGATGAGAGCAAGAGGAC 300
DB 241 CCCACACCCCGGAGCCAGAGGAGATCTTGGGATCAGATGAGAGCAAGAGGAC 300
QY 301 CTTGGGACTACTGCAAGGTGGATATCCAGTGAAATTTGGAGACCTCTTCAATGGC 360
DB 301 CTTGGGACTACTGCAAGGTGGATATCCAGTGAAATTTGGAGACCTCTTCAATGGC 360
QY 361 CGGTATCATGTTATTAGAAAGCTTGGATGGGGCACTTCTTACTGTCTGGCTGTGCTGG 420
DB 361 CGGTATCATGTTATTAGAAAGCTTGGATGGGGCACTTCTTACTGTCTGGCTGTGCTGG 420

QY 421 GATATGCGGGGAAAGATTTTGTGCAATGAAAGTTGTAAAAAGTCCCGAGCATTATACG 480
DB 421 GATATGCGGGGAAAGATTTTGTGCAATGAAAGTTGTAAAAAGTCCCGAGCATTATACG 480
QY 481 GAGACAGCCCTTGGATGAAATAAAATTTGCTCAAAATGTGTTTCGAGAAAGTATCCAGTGAC 540
DB 481 GAGACAGCCCTTGGATGAAATAAAATTTGCTCAAAATGTGTTTCGAGAAAGTATCCAGTGAC 540
QY 541 CCAAAACAAAGACATGGTGGTCCAGCTCATTTGACGACTTCAAGATTTTCAAGATTTTCAAGTGGG 600
DB 541 CCAAAACAAAGACATGGTGGTCCAGCTCATTTGACGACTTCAAGATTTTCAAGATTTTCAAGTGGG 600
QY 601 ATACATGCTGCTCATGCTTTCGAAAGTACTTGGCCACCCTCTCTCAAGTGGATCATCAAA 660
DB 601 ATACATGCTGCTCATGCTTTCGAAAGTACTTGGCCACCCTCTCTCAAGTGGATCATCAAA 660
QY 661 TCCAACTATCAAGGGCTCCAGTACGTTGTGTGAAGAGTATCATTTCCAGAGTCTCTTCAA 720
DB 661 TCCAACTATCAAGGGCTCCAGTACGTTGTGTGAAGAGTATCATTTCCAGAGTCTCTTCAA 720
QY 721 GGGTTAGATTACTTACACAGTAAAGTGCAGATCATTTCTACTGACATAAAGCCGGAAT 780
DB 721 GGGTTAGATTACTTACACAGTAAAGTGCAGATCATTTCTACTGACATAAAGCCGGAAT 780
QY 781 ATCTTGTATGTTGTGGATGATGCATATGTGAGAAGATGGCAGCTGAGGCCACTGAGTGG 840
DB 781 ATCTTGTATGTTGTGGATGATGCATATGTGAGAAGATGGCAGCTGAGGCCACTGAGTGG 840
QY 841 CAGAAAGCAGGTGCTCTCTCTCTTTCAGGGTCTGAGTACGGCTCCACAGCAGAAA 900
DB 841 CAGAAAGCAGGTGCTCTCTCTCTTTCAGGGTCTGAGTACGGCTCCACAGCAGAAA 900
QY 901 CCTATAGGAAAAATATCTAAAAACAAAAAGAAAAAATCTGAAAAAGAAAAGAGAGGCGAG 960
DB 901 CCTATAGGAAAAATATCTAAAAACAAAAAGAAAAAATCTGAAAAAGAAAAGAGAGGCGAG 960
QY 961 GCTGAGTTATTTGGAGAGCGCTGAGGAGTAGAAGATTTGGAGCGAGAACTGGAAGG 1020
DB 961 GCTGAGTTATTTGGAGAGCGCTGAGGAGTAGAAGATTTGGAGCGAGAACTGGAAGG 1020
QY 1021 AAAATAATAGAAAAAATCATCACCTCAGCTGCACCTTCCAAATGACAGAGTGGCGAATAC 1080
DB 1021 AAAATAATAGAAAAAATCATCACCTCAGCTGCACCTTCCAAATGACAGAGTGGCGAATAC 1080
QY 1081 TGCCCGAGGTTGAAACTAAAAACAACAGGATTTAGAGGAGCGCTGAGGCGAGACTGCA 1140
DB 1081 TGCCCGAGGTTGAAACTAAAAACAACAGGATTTAGAGGAGCGCTGAGGCGAGACTGCA 1140
QY 1141 AAGGCAATTTGGTGAAGCTGAGGACAGGAGAAAGAGATGCTGAGAAAGAAACATTT 1200
DB 1141 AAGGCAATTTGGTGAAGCTGAGGACAGGAGAAAGAGATGCTGAGAAAGAAACATTT 1200
QY 1201 GAAAAAGATGAAGATGATGATAGTACAGGAACTTTCGAACTAGACCTACCTGGATAGAA 1260
DB 1201 GAAAAAGATGAAGATGATGATAGTACAGGAACTTTCGAACTAGACCTACCTGGATAGAA 1260
QY 1261 TCACCTAAAAACCAATGGCCATATTGAGAAATGGCCATTTCTCACTGAGCAGCAACTGGAC 1320
DB 1261 TCACCTAAAAACCAATGGCCATATTGAGAAATGGCCATTTCTCACTGAGCAGCAACTGGAC 1320
QY 1321 GATGAAGATGATGATGAAGAGACTGCCCCAAATCTCTGAGGAATATATCTTGTATGAGCCA 1380
DB 1321 GATGAAGATGATGATGAAGAGACTGCCCCAAATCTCTGAGGAATATATCTTGTATGAGCCA 1380
QY 1381 AATGAGAAAGTATTACACATATAGCAGCTCTTATGAACAATTTCAATGGTGAATGGCCA 1440
DB 1381 AATGAGAAAGTATTACACATATAGCAGCTCTTATGAACAATTTCAATGGTGAATGGCCA 1440
QY 1441 AATGAGCAATATAAAATTTCCGAGTCCAGAGTCTTCCACCTCGTTGTCTCT 1500
DB 1441 AATGAGCAATATAAAATTTCCGAGTCCAGAGTCTTCCACCTCGTTGTCTCT 1500
QY 1501 GGATCCTTAGAAACCTGTGGCCTGTGGCTCTGTCTTCTGAGGGATCACCACTTACTGAG 1560

Qy	153	TCCGAAAAAGCGGAGCCTCAACAGAAAGCTCTCTTTAGTTTCTCTCTCTCTCAACCGCCACC	212
Db	208	TTCAAGAAAGCGCGAGCCTCAACAGAAAGCTCTTTAGTTTCTCTCTCTCTCAACCGCCACC	267
Qy	213	ACCACCACCGCGCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGGAGATCCT	272
Db	268	ACCACCACCGCGCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGGAGATCCT	327
Qy	273	GGGATCAGATGATGAGGAGCAAGAGGCCCTTCGGGACTACTGCAAAAGTGGATATCATCC	332
Db	328	GGGATCAGATGATGAGGAGCAAGAGGCCCTTCGGGACTACTGCAAAAGTGGATATCATCC	387
Qy	333	AGTGAAAAATGGAGACCTCTTCAATGGCGGTATCATGTTATTAGAAAGCTTGGATGGG	392
Db	388	AGTGAAAAATGGAGACCTCTTCAATGGCGGTATCATGTTATTAGAAAGCTTGGATGGG	447
Qy	393	GCACCTTCTACTGTCTGTGCTGTCTGGGATATGACGGGGAAGATTTGTTGCAATGAA	452
Db	448	GCACCTTCTACTGTCTGTGCTGTCTGGGATATGACGGGGAAGATTTGTTGCAATGAA	507
Qy	453	AGTTGTAATAAGTCCCGACGATTAACGGAGACAGCCCTTGGATGAAATAAAATTCGCTCAA	512
Db	508	AGTTGTAATAAGTCCCGACGATTAACGGAGACAGCCCTTGGATGAAATAAAATTCGCTCAA	567
Qy	513	ATGTGTTGAGAAAAGTGATCCAGTGACCCAAAACAAAGACATGTGTGTTCAGCTCATTTGA	572
Db	568	ATGTGTTGAGAAAAGTGATCCAGTGACCCAAAACAAAGACATGTGTGTTCAGCTCATTTGA	627
Qy	573	CGACTTCAAGATTTTCAGCATGAATGGGATACATGTCTGCAATGATCTTCGAAGTACTTTGG	632
Db	628	CGACTTCAAGATTTTCAGCATGAATGGGATACATGTCTGCAATGATCTTCGAAGTACTTTGG	687
Qy	633	CCACCATCTCCTCAAGTGGATCATCAAAATCCAACCTATCAAGGCCCTCCAGTACGCTTGTGT	692
Db	688	CCACCATCTCCTCAAGTGGATCATCAAAATCCAACCTATCAAGGCCCTCCAGTACGCTTGTGT	747
Qy	693	GAAGAGTATCAATTCGACAGGTCCTTCAAGGGTTAGATTATCTTACACAGTAAAGTCAAGAT	752
Db	748	GAAGAGTATCAATTCGACAGGTCCTTCAAGGGTTAGATTATCTTACACAGTAAAGTCAAGAT	807
Qy	753	CATTTCATCTGACATAAGCCGGAATATCTTCGATGTGTGTGATGATGCATATGTGAG	812
Db	808	CATTTCATCTGACATAAGCCGGAATATCTTCGATGTGTGTGATGATGCATATGTGAG	867
Qy	813	AAGAATGGCAGCTGAGGCGCACTGAGTGGCAGAAAGCAGGTGCTCCTCTCTTCAGGGTC	872
Db	868	AAGAATGGCAGCTGAGGCGCACTGAGTGGCAGAAAGCAGGTGCTCCTCTCTTCAGGGTC	927
Qy	873	TGCAGTGAGTACGGCTCCACAGCAGAAAACCTATAGGAAAAATATCTAAAAACAAAAAGAA	932
Db	928	TGCAGTGAGTACGGCTCCACAGCAGAAAACCTATAGGAAAAATATCTAAAAACAAAAAGAA	987
Qy	933	AAAACTGAAAAAGAAAACGAAAGCGAGGCTGAGTTATTGGAGAAAGCGCCTGCAGGAGAT	992
Db	988	AAAACTGAAAAAGAAAACGAAAGCGAGGCTGAGTTATTGGAGAAAGCGCCTGCAGGAGAT	1047
Qy	993	AGAAAGATTGGAGCGGAAGCTGAAAGGAAAAATTAATAGAGAAAAACATCACTCAGCTGC	1052
Db	1048	AGAAAGATTGGAGCGGAAGCTGAAAGGAAAAATTAATAGAGAAAAACATCACTCAGCTGC	1107
Qy	1053	ACCTTCCAATGACCAGGATGCGAATATCTGCCAGAGGTGAAACTTAAAAACAAACAGGATT	1112
Db	1108	ACCTTCCAATGACCAGGATGCGAATATCTGCCAGAGGTGAAACTTAAAAACAAACAGGATT	1167
Qy	1113	AGAGGAGCGGCTGAGGCGAGACTGCAAGGACAAATGGTGAAGCTGAGGACGAGGAAGA	1172
Db	1168	AGAGGAGCGGCTGAGGCGAGACTGCAAGGACAAATGGTGAAGCTGAGGACGAGGAAGA	1227
Qy	1173	GAAGAAGATGCTCGAGAAAGAAAACATTGAAAAAGATGAAGATGATGTAGATCAGGAAC	1232
Db	1228	GAAGAAGATGCTCGAGAAAGAAAACATTGAAAAAGATGAAGATGATGTAGATCAGGAAC	1287
Qy	1233	TGCGAAACATAGACCCCTACGTGGATAGAAATCACCTTAAAAACCAATGGCCATATTGAGAAATGG	1292

[illegible]

Db 2368 GCTTTGAGTACTGTGTTTGACATTTTGTGCTGCTTTTGTGCACTGTGATCCTCGGGGAAGGGT 2427
Qy 2373 AGTCTTTTGTCTTCACTAAGTCTAGTTTACTGACCAATTTTCTTCTGGAACAATAACATGT 2432
Db 2428 AGTCTTTTGTCTTCACTAAGTCTAGTTTACTGACCAATTTTCTTCTGGAACAATAACATGT 2487
Qy 2433 CTCTAAGCAATGTTTCTTGTGTTGTGTCATTTCAAATGTCATTTTGTGAATGAAATAAT 2492
Db 2488 CTCTAAGCAATGTTTCTTGTGTTGTGTCATTTCAAATGTCATTTTGTGAATGAAATAAT 2547
Qy 2493 ACTTTCCCTTTGTGTTTGGCAGGTTTGTGTAATCTATTATGAAGAAATATTTAGCTGA 2552
Db 2548 ACTTTCCCTTTGTGTTTGGCAGGTTTGTGTAATCTATTATGAAGAAATATTTAGCTGA 2607
Qy 2553 GTACTATATAATTAACAATCTTAAGAAATATCAAGTTTGGGAACCAAGAAATAATAGCAAGG 2612
Db 2608 GTACTATATAATTAACAATCTTAAGAAATATCAAGTTTGGGAACCAAGAAATAATAGCAAG 2664
Qy 2613 GAAATGTACAAATTTTATCTTCTGCAAAAGGACATCATTTCTCTGTATTATAGTGTATGTA 2672
Db 2665 GAAATGTACAAATTTTATCTTCTGCAAAAGGACATCATTTCTCTGTATTATAGTGTATGTA 2724
Qy 2673 ATGCACCTGTAAATTTACTTTTGGATTAATAATATGGAGGGGGAGCTCAAAATTCAGAAA 2732
Db 2725 ATGCACCTGTAAATTTACTTTTCCATTAATAATATGGGA-GGGGGACTCAAAATTTTCAGAAA 2783
Qy 2733 AGCTAAAAA 2741
Db 2784 AGCTACCAA 2792

RESULT 5

US-10-198-846-10286
; Sequence 10286, Application US/10198846
; Publication No. US20030099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Streimann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10286
; LENGTH: 4698
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-10286

Query Match 77.9%; Score 2533.8; DB 14; Length 4698;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2568; Conservative 0; Mismatches 17; Indels 4; Gaps 2;
Qy 153 TCCGAAAAACCGGAGCCTCAACAGAAAGCTCTTTAGTTTCCTCTCCCTCCACCGCCACC 212
Db 254 TTCAGAAAGCGGAGCCTCAACAGAAAGCTCTTTAGTTTCCTCTCCCTCCACCGCCACC 313
Qy 213 ACCACCAACCGGAGCCTTTCAGACCCCAACCCCGGAGCAGAGAGAGATCCT 272
Db 314 ACCACCAACCGGAGCCTTTCAGACCCCAACCCCGGAGCAGAGAGAGATCCT 373
Qy 273 GGGATCAGATGATGAGGAGCAAGACCTCTGGGACTACTGCAAAAGGTGGATATCATCC 332
Db 374 GGGATCAGATGATGAGGAGCAAGACCTCTGGGACTACTGCAAAAGGTGGATATCATCC 433
Qy 333 AGTAAAAATTGGAGACCTCTTCAATGCGCGGTATCATGTTATTAGAAAGCTTGATGGGG 392

Db 434 AGTCAAAATTTGGAGACCTCTTCAATGCGCGGTATCATGTTATTAGAAAGCTTGATGGGG 493
Qy 393 GCATCTCTCTACTGTCTGGCTGTCTGGGATATCGAGGGGAAAGATTTTCTTCAATGAA 452
Db 494 GCATCTCTCTACTGTCTGGCTGTCTGGGATATCGAGGGGAAAGATTTTCTTCAATGAA 553
Qy 453 AGTTGTAAGATGCTCCAGCATTTATACGGAGACAGCTTGGATGAATAAATAATGCTCAA 512
Db 554 AGTTGTAAGATGCTCCAGCATTTATACGGAGACAGCTTGGATGAATAAATAATGCTCAA 613
Qy 513 ATGTGTTTCGAAAGATGATCCAGTGACCCCAACCAAGACATGCTGTCTCCAGCTCATTTGA 572
Db 614 ATGTGTTTCGAAAGATGATCCAGTGACCCCAACCAAGACATGCTGTCTCCAGCTCATTTGA 673
Qy 573 CGATCTCAAGATTTCAAGGATGAATGGGATACATGCTGTGATGCTTCTCGAAGTACTTGG 632
Db 674 CGATCTCAAGATTTCAAGGATGAATGGGATACATGCTGTGATGCTTCTCGAAGTACTTGG 733
Qy 633 CCACATCTCTCTCAAGTGGATCAATCAATCAATCAATCAAGGCTCCAGTACGTTGTGT 692
Db 734 CCACATCTCTCTCAAGTGGATCAATCAATCAATCAATCAAGGCTCCAGTACGTTGTGT 793
Qy 693 GAAAGATCATCTTCGACAGGCTCTTCAAGGTTAGATTACTTACACAGTAAAGTCAAGAT 752
Db 794 GAAAGATCATCTTCGACAGGCTCTTCAAGGTTAGATTACTTACACAGTAAAGTCAAGAT 853
Qy 753 CATTCATACCTGACATAAAGCGGAAATAATCTTGTGATGCTGTGGATGATCATATGTGAG 812
Db 854 CATTCATACCTGACATAAAGCGGAAATAATCTTGTGATGCTGTGGATGATCATATGTGAG 913
Qy 813 AAGAAATGGAGCTGAGGCTCTGAGTGGCAAGAGAGGCTCTCTCTCTCTCAGGGTC 872
Db 914 AAGAAATGGAGCTGAGGCTCTGAGTGGCAAGAGAGGCTCTCTCTCTCTCAGGGTC 973
Qy 932 TGCAGTGTAGTACGCTCCACAGAGAAACCTTATAGGAAATAATCTTAAAAACAAGAA 932
Db 974 TGCAGTGTAGTACGCTCCACAGAGAAACCTTATAGGAAATAATCTTAAAAACAAGAA 1033
Qy 933 AAAAATGTAAGAAACAGAGAGGCTGTGATTTATGGAGAGGCTCTGAGGAGAT 992
Db 1034 AAAAATGTAAGAAACAGAGAGGCTGTGATTTATGGAGAGGCTCTGAGGAGAT 1093
Qy 993 AGAAGATTTGAGCGAGAGCTGAAAGAAATATAGAGAAACATCACTCAGCTGC 1052
Db 1094 AGAAGATTTGAGCGAGAGCTGAAAGAAATATAGAGAAACATCACTCAGCTGC 1153
Qy 1053 ACCTTCCAAATGACAGGATGCGAATCTGCCAGAGGTGAAACTTAAAAACAACAGGAT 1112
Db 1154 ACCTTCCAAATGACAGGATGCGAATCTGCCAGAGGTGAAACTTAAAAACAACAGGAT 1213
Qy 1172 AGAGAGGCGGTGAGGCGAGAGCTGCAAGGACAAATGCTGAAGTGAAGGACAGGAGAG 1172
Db 1214 AGAGAGGCGGTGAGGCGAGAGCTGCAAGGACAAATGCTGAAGTGAAGGACAGGAGAG 1273
Qy 1232 GAAAGAGATGCTGAAAGAAACATTTGAAAGATGCAAGATGATGATGATGATGATGATGAT 1232
Db 1274 GAAAGAGATGCTGAAAGAAACATTTGAAAGATGCAAGATGATGATGATGATGATGATGAT 1333
Qy 1292 TCGCAACATAGACCTTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1292
Db 1334 TCGCAACATAGACCTTACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1393
Qy 1352 CCCATTTCTCACTGGAGCGCAACTGGAGCGATGATGATGATGATGATGATGATGATGATGAT 1352
Db 1394 CCCATTTCTCACTGGAGCGCAACTGGAGCGATGATGATGATGATGATGATGATGATGATGAT 1453
Qy 1412 TCCTGAGGATATATCTTGTGAGCCAAATGCAAAAGTATTACATATAGCAGCTC 1412
Db 1454 TCCTGAGGATATATCTTGTGAGCCAAATGCAAAAGTATTACATATAGCAGCTC 1513
Qy 1413 CTATGAAACAATTTCAATGTTGATTTGCCAAATGGACGACATAAAATTTCCGAGTCAAGTT 1472

[illegible]

RESULT 6

US-10-425-114-16424

US-10-423-114-16424
: Sequence 16424. Application US/10425114

; sequence 16424, Application US/ : Publication No. IIS20040034888A1

; Publication No: US20
: GENERAL INFORMATION:

APPLICANT: Liu Jindong

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

; APPLICANT: Screen, Steve

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

; TITLE OF INVENTION: Nucleic Acid Molecule

; TITLE OF INVENTION: Plants and

; FILE REFERENCE

; CURRENT APPLIC

1. **CURRENT FI**

; NUMBER OF SEQ ID NOS: 7

; SEQ ID NO 1

; LENGTH: 3

TYPE: DNA

ORGANISM: Homo sapiens

; ORGANISM: Homo sapiens
; FEATURE:

OTHER INFORMATION: Clone ID: LIB3063-134-E5 FLI

OTHER INFORMATION: CIONE ID: DIB3003-134-ES_FDI
US-10-425-114-16424

42401-4TY-CZF-AT-SO

Query Match 77.7%: Score 2529: DB 16: Length 3699:

Query Match 77.7%; Score 2329; DB 16; Length 3699;
Best Local Similarity 99.1%; Pred NC 0.

BEST LOCAL SIMILARITY	99.18;	PRED. NO. 0;
MATCHES 2555.	CONSERVATIVE	0.
MISMATCHES	20.	INDOLS 4.

Matches 2565; Conservative 0; Mismatches 20; Indels 4; Gaps 2

153

QY 153 TCCGAAAAGCCGGAGCCTCAACAGAAAGCTCCTCTCTCCACGCCACC 212

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

D_b 199 TTCAGAAAGGCCGGAGCCTCAACAGAAAGCTCCTTAGTTCTCCTCCTTACCGCCACC 258

4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
8

QY
213 ACCACCACCGCCACCTTTGCCAGACCCACACCCCCGGAGCCAGAGGAGATCCT 272

Age Group	Total	Male	Female	Male	Female
18-24	15%	12%	18%	10%	20%
25-34	25%	22%	28%	18%	30%
35-44	30%	28%	32%	22%	35%
45-54	20%	18%	22%	15%	25%
55-64	15%	12%	18%	10%	20%
65-74	10%	8%	12%	6%	14%
75+	5%	4%	6%	3%	7%

Db 559 ATGTGTTCCGAGAAAGTGATCCCGAGTGACCCAAACAAAGACATGTTGGTCCAGCTCAITGA 618
QY 573 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTCTGCAATGCTCTCCAGGACTTGG 632
Db 619 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTCTGCAATGCTCTCCAGGACTTGG 678
QY 633 CCACCATCTCTCAAGTGGATCATCAAAATCAAACTATCAAGGCCTCCAGGACTTGGTGTGT 692
Db 679 CCACCATCTCTCAAGTGGATCATCAAAATCAAACTATCAAGGCCTCCAGGACTTGGTGTGT 738
QY 693 GAAGAGTATCAITTCGACAGGCTCTTCAAGGGTTAGATTACTTACACAGTAAGTGAAGAT 752
Db 739 GAAGAGTATCAITTCGACAGGCTCTTCAAGGGTTAGATTACTTACACAGTAAGTGAAGAT 798
QY 753 CATTCATCTGACATATAAGCCGGAAATATCTTGTATGTTGTGGATGATGATATGTCAG 812
Db 799 CATTCATCTGACATATAAGCCGGAAATATCTTGTATGTTGTGGATGATGATATGTCAG 858
QY 813 AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAACAGGCTGCTCTCTCTCTTCCAGGGTC 872
Db 859 AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAACAGGCTGCTCTCTCTCTTCCAGGGTC 918
QY 873 TGCAGTGTAGTACGGCTCCACAGCAAAACCTATAGGAAATATCTTAAACAAAGAA 932
Db 919 TGCAGTGTAGTACGGCTCCACAGCAAAACCTATAGGAAATATCTTAAACAAAGAA 978
QY 933 AAAACTGAAAAGAAACAGAGGCGAGCTGAGTTATTTGGAGAGCGCTCCAGGAGAT 992
Db 979 AAAACTGAAAAGAAACAGAGGCGAGCTGAGTTATTTGGAGAGCGCTCCAGGAGAT 1038
QY 993 AGAAGAAATTTGGACGAGAAAGCTGAAAGGAAATATAGAGAAACATCACTCAGCTGC 1052
Db 1039 AGAAGAAATTTGGACGAGAAAGCTGAAAGGAAATATAGAGAAACATCACTCAGCTGC 1098
QY 1053 ACCTTCCAATGACCGAGATGGCGAATACCTGCCAGAGGTGAACTTAAACAAAGAGAT 1112
Db 1099 ACCTTCCAATGACCGAGATGGCGAATACCTGCCAGAGGTGAACTTAAACAAAGAGAT 1158
QY 1113 AGAGGAGGCGCTGAGCGAGAGACTGCAAGAGCAATGCTGAAGCTGAGGACGAGCAAGA 1172
Db 1159 AGAGGAGGCGCTGAGCGAGAGACTGCAAGAGCAATGCTGAAGCTGAGGACGAGCAAGA 1218
QY 1173 GAAAGAGATGCTGAGAGAAAGAAACATTTGAAAGATGAAGATGATGTAGATCAGGAAT 1232
Db 1219 GAAAGAGATGCTGAGAGAAAGAAACATTTGAAAGATGAAGATGATGTAGATCAGGAAT 1278
QY 1233 TGGCAACATAGACCTTACGTGGATAGAACTCACTTAAACCAATGGCCATATTTGAGAAATGG 1292
Db 1279 TGGCAACATAGACCTTACGTGGATAGAACTCACTTAAACCAATGGCCATATTTGAGAAATGG 1338
QY 1293 CCCATCTCACTGGAGCAGCACTGGAAGATGAAGATGATGAAGAGAGACTGCCCAAA 1352
Db 1339 CCCATCTCACTGGAGCAGCACTGGAAGATGAAGATGATGAAGAGAGACTGCCCAAA 1398
QY 1353 TCCTGAGGAATAAATCTTGATGAGCCAAATGAGCAAAAGTGAATACATATAGCAGCTC 1412
Db 1399 TCCTGAGGAATAAATCTTGATGAGCCAAATGAGCAAAAGTGAATACATATAGCAGCTC 1458
QY 1413 CTATGACAAATTCATGTTGAATTTGCCAATGACGACATATAAATTTCCGAGTCAAGTT 1472
Db 1459 CTATGACAAATTCATGTTGAATTTGCCAATGACGACATATAAATTTCCGAGTCAAGTT 1518
QY 1473 CCCAGAGTTTCCACCTCGTGTCTCTGATCTCTGATCTCTTGAACCTGTGGCTCGCGCTCTGT 1532
Db 1519 CCCAGAGTTTCCACCTCGTGTCTCTGATCTCTGATCTCTTGAACCTGTGGCTCGCGCTCTGT 1578
QY 1533 GCTTTCGAGGATACCACTTACTGAGCAAGAGGAGCAGCTCCATCCCATGACAGAG 1592
Db 1579 GCTTTCGAGGATACCACTTACTGAGCAAGAGGAGCAGCTCCATCCCATGACAGAG 1638
QY 1593 CAGAACGGTTTCAGCCTCCAGTACTGGGGATTTGCCAAAGCAAAACCCGGGAGCTGA 1652
Db 1639 CAGAACGGTTTCAGCCTCCAGTACTGGGGATTTGCCAAAGCAAAACCCGGGAGCTGA 1698

QY 1653 CTTGTTGGTGAATCCCTCGATCCGCGAATGCGATATAAATTTAGAGTAAAAATTTGCTGA 1712
Db 1699 CTTGTTGGTGAATCCCTCGATCCGCGAATGCGATATAAATTTAGAGTAAAAATTTGCTGA 1758
QY 1713 CTTGGGAAATCTGTTGGTGCATAAACAATTTCAAGGAGACATCCAGACGCTCAGTA 1772
Db 1759 CTTGGGAAATCTGTTGGTGCATAAACAATTTCAAGGAGACATCCAGACGCTCAGTA 1818
QY 1773 CCGCTCATAGAGTGTATTAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGGAGCAC 1832
Db 1819 CCGCTCATAGAGTGTATTAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGGAGCAC 1878
QY 1833 GGGCTGTATGGCAATTTGAGCTGGCAACGCGAGATTAATTTGTTTGAACCAATCTCTGGGGA 1892
Db 1879 GGGCTGTATGGCAATTTGAGCTGGCAACGCGAGATTAATTTGTTTGAACCAATCTCTGGGGA 1938
QY 1893 AGACTATTCAGAGACGAGACACATAGCCCAATCATAGAGCTGTAGGAGTATTC 1952
Db 1939 AGACTATTCAGAGACGAGACACATAGCCCAATCATAGAGCTGTAGGAGTATTC 1998
QY 1953 AAGCACTTTGCTCTATCTGAAAAATATTTCTCGGAAATTTCTCAATCGCAGAGAGAACT 2012
Db 1999 AAGCACTTTGCTCTATCTGAAAAATATTTCTCGGAAATTTCTCAATCGCAGAGAGAACT 2058
QY 2013 GCGACATCAACCAAGCTGAAAGCCCTGGAGCCTCTTTGATGTACTTTGAAAAAGTATGG 2072
Db 2059 GCGACATCAACCAAGCTGAAAGCCCTGGAGCCTTTTGTGATGTACTTTGAAAAAGTATGG 2118
QY 2073 CTGGCCCATCAAGATGCTGCACAGTTTACAGATTTCTCTGATCCGATGTTTGAAGATGGT 2132
Db 2119 CTGGCCCATCAAGATGCTGCACAGTTTACAGATTTCTCTGATCCGATGTTTGAAGATGGT 2178
QY 2133 TCCAGAAAAACGAGCCTCAGCTGGCGAATGCTTTGGCAATGCTTTGGTGAATTTCTTAGCA 2192
Db 2179 TCCAGAAAAACGAGCCTCAGCTGGCGAATGCTTTGGCAATGCTTTGGTGAATTTCTTAGCA 2238
QY 2193 AATTCACCAATATGCAATCTGAGCTAGCAAAATGTTCCAGATACATTTGGACCTTAAACGG 2252
Db 2239 AATTCACCAATATGCAATCTGAGCTAGCAAAATGTTCCAGATACATTTGGACCTTAAACGG 2298
QY 2253 TGACTCTCATTTTAAACAGGATTAACAGTGGCTTCAAGTGGCTTCACTCTCAGACCTTTATTTT 2312
Db 2299 TGACTCTCATTTTAAACAGGATTAACAGTGGCTTCAAGTGGCTTCACTCTCAGACCTTTATTTT 2358
QY 2313 GCTTTGAGGTACTGTTGTTGCAATTTTGTCTTTTGTGCACTGTGATCTCTGGGAGGGT 2372
Db 2359 GCTTTGAGGTACTGTTGTTGCAATTTTGTCTTTTGTGCACTGTGATCTCTGGGAGGGT 2418
QY 2373 AGTCTTTTGTCTTCAAGTAACTGTTTACTGACCAATTTTCTCTGGAAACAAATTAACATGT 2432
Db 2419 AGTCTTTTGTCTTCAAGTAACTGTTTACTGACCAATTTTCTCTGGAAACAAATTAACATGT 2478
QY 2433 CTCTAAGCATTTCTTGTGTGTGTGATGACATTTCAAATGTCATTTTGTGAATGAAAAAT 2492
Db 2479 CTCTAAGCATTTCTTGTGTGTGTGATGACATTTCAAATGTCATTTTGTGAATGAAAAAT 2538
QY 2493 ACTTTCCCTTTGTTTTGGCAGGTTTGTGTAACCTATTTTGAAGAAATATTTTACTGTA 2552
Db 2539 ACTTTCCCTTTGTTTTGGCAGGTTTGTGTAACCTATTTTGAAGAAATATTTTACTGTA 2598
QY 2553 GTACTATATAATTTTAAATCTTAAGAAATTTCAAGTTTGAACCAAGAAATA --- GCAAG 2612
Db 2599 GTACTATATAATTTTAAATCTTAAGAAATTTCAAGTTTGAACCAAGAAATA --- GCAAG 2655
QY 2613 GAAATGTACAAATTTTATCTTGGCAAGGGGACATCATCTCTGTATTTATGTGTATGTAA 2672
Db 2656 GAAATGTACAAATTTTATCTTGGCAAGGGGACATCATCTCTGTATTTATGTGTATGTAA 2715
QY 2673 ATGCAACCTGTAAATGTTACTTTTGAATTAATGAGGGGGGAGCTCAAAATTTTCAAGAAA 2732
Db 2716 ATGCAACCTGTAAATGTTACTTTTCCATTTAAATATATGGGA - GGGGGAGCTCAAAATTTTCAAGAAA 2774

Qy	2733	AGCTAAAAA	2741
Db	2775	AGTACCAA	2783
RESULT 7			
US-09-981-397A-21			
; Sequence 21, Application US/09981397A			
; Publication No. US20030082519A1			
; GENERAL INFORMATION:			
; APPLICANT: Axxima Pharmaceuticals AG			
; APPLICANT: Schubart, Daniel			
; APPLICANT: Habenberger, Peter			
; APPLICANT: Stein-Gerlach, Matthias			
; APPLICANT: Bevec, Dorian			
; TITLE OF INVENTION: Cellular Kinases Involved in Cytomegalovirus Infection and their			
; TITLE OF INVENTION: Inhibition			
; FILE REFERENCE: AXM-004.1 US			
; CURRENT APPLICATION NUMBER: US/09/981.397A			
; CURRENT FILING DATE: 2002-06-28			
; PRIOR APPLICATION NUMBER: 60/240,750			
; PRIOR FILING DATE: 2000-10-16			
; NUMBER OF SEQ ID NOS: 22			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 21			
; LENGTH: 3745			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-09-981-397A-21			
Query Match 76.8%; Score 2497; DB 10; Length 3745;			
Best Local Similarity 98.8%; Pred. No. 0;			
Matches 2559; Conservative 0; Mismatches 20; Indels 10; Gaps 4;			
Qy	153	TCCGAAAAGCGCGAGCCTCAACAGAAAGTCCTTTAGTTCTCTCTCTCCACGCCACC	212
Db	217	TTCAGAAAGCGCGAGCCTCAACAGAAAGTCCTTTAGTTCTCTCTCTCCACGCCACC	276
Qy	213	ACCACCACCGGCACCTTTGGCCAGACCCACACCCCAGGAGCCAGAGGAGATCCT	272
Db	277	ACCACCAACCGGCACCTTTGGCCAGACCCACACCCCAGGAGCCAGAGGAGATCCT	336
Qy	273	GGGATCATGATGAGGAGCAAGAGGACCTTCGCGACTACTGCAAGGTGGATATCATCC	332
Db	337	GGGATCATGATGAGGAGCAAGAGGACCTTCGCGACTACTGCAAGGTGGATATCATCC	396
Qy	333	AGTGAAAATTGGAGACCTCTTCAATGCCCGTATCATGTTATTAGAAGCTTGGATGGG	392
Db	397	AGTGAAAATTGGAGACCTCTTCAATGCCCGTATCATGTTATTAGAAGCTTGGATGGG	456
Qy	393	GCATCTCTACTGCTGTGGCTGTCTGGGATATCGAGGGGAAAAGATTGTTGCAATGA	452
Db	457	GCATCTCTACTGCTGTGGCTGTCTGGGATATCGAGGGGAAAAGATTGTTGCAATGA	516
Qy	453	AGTTGTAATAAGTCCCGACGATTTATACGAGACAGCCTTGGATGAAAATAAAATTGCTCAA	512
Db	517	AGTTGTAATAAGTCCCGACGATTTATACGAGACAGCCTTGGATGAAAATAAAATTGCTCAA	576
Qy	513	ATGTGTTTCGAGAAAGTAGTCCAGTGAGCCCCAAAACAAGACATGGTGGTCCAGCTCATGA	572
Db	577	ATGTGTTTCGAGAAAGTAGTCCAGTGAGCCCCAAAACAAGACATGGTGGTCCAGCTCATGA	636
Qy	573	CGACTTCAAGATTTTCAGCATGAATGGATACATGCTGCGATGGTCTTCGAAGTACTTGG	632
Db	637	CGACTTCAAGATTTTCAGCATGAATGGATACATGCTGCGATGGTCTTCGAAGTACTTGG	696
Qy	633	CCACCATCTCTCAAGTGGATCATCAAACTATCAAGGCCCTCCAGTACGTTGTGT	692
Db	697	CCACCATCTCTCAAGTGGATCATCAAACTATCAAGGCCCTCCAGTACGTTGTGT	756
Qy	693	GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTATACACAGTAAGTCAAGAT	752
Db	757	GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTATACACAGTAAGTCAAGAT	816

QY	1833	GGCGTGATGGCATTTGAGCTGCGAAACGGGAGATTATTTGTTGAACCAATTCCTGGGGA	1892
DB	1894	GGCGTGATGGCATTTGAGCTGCGAAACGGGAGATTATTTGTTGAACCAATTCCTGGGGA	1953
QY	1893	AGACTATTCAGAGACGAAGACACATAGCCACATCATAGAGCTGCTTAGCGAGTATTC	1952
DB	1954	AGACTATTCAGAGACGAAGACACATAGCCACATCATAGAGCTGCTTAGCGAGTATTC	2013
QY	1953	AAGGCATTTGCTCTATCTCGAAAAATATTTCTCGGAAATTTCTTCAATCGCAGAGGAACT	2012
DB	2014	AAGGCATTTGCTCTATCTCGAAAAATATTTCTCGGAAATTTCTTCAATCGCAGAGGAACT	2073
QY	2013	GGGACACATCAACAAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGGAAGATGCG	2072
DB	2074	GGGACACATCAACAAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGGAAGATGCG	2133
QY	2073	CTGGCCCGATGAAGATGCTGGACAGTTTACAGATTTCTTGATCCCGATGTTAGAAATGCT	2132
DB	2134	CTGGCCCGATGAAGATGCTGGACAGTTTACAGATTTCTTGATCCCGATGTTAGAAATGCT	2193
QY	2133	TCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCCTTGGGCATCCTTGGTGTGAATTCCTTAGCA	2192
DB	2194	TCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCCTTGGGCATCCTTGGTGTGAATTCCTTAGCA	2250
QY	2193	AATTTCTACCAATATTTGCAATTTCTGAGCTAGCAAAATGTTCCAGTACATTTGAGACCTTAAACGG	2252
DB	2251	AATTTCTACCAATATTTGCAATTTCTGAGCTAGCAAAATGTTCCAGTACATTTGAGACCTTAAACGG	2310
QY	2253	TGACTCTCATTTTAAACAGATTTAACAAGTGAGCTGGCTTCATCCCTCAGACCTTTATTTT	2312
DB	2311	TGACTCTCATTTTAAACAGATTTAACAAGTGAGCTGGCTTCATCCCTCAGACCTTTATTTT	2370
QY	2313	GCTTTGAGGTACTGTGTTTGACATTTTGTGCACTGTGATCCTTGGGGAAGGCT	2372
DB	2371	GCTTTGAGGTACTGTGTTTGACATTTTGTGCACTGTGATCCTTGGGGAAGGCT	2430
QY	2373	AGTCTTTTGTCTTCAGCTAAGTAGTTTACGACCATTTTCTTCTGGAACAATTAACATGT	2432
DB	2431	AGTCTTTTGTCTTCAGCTAAGTAGTTTACGACCATTTTCTTCTGGAACAATTAACATGT	2490
QY	2433	CTCTAAGCATTTGTTCTTGTGTTGTCGACATTCABAATGTCAATTTTTTTCGAATGAANAAT	2492
DB	2491	CTCTAAGCATTTGTTCTTGTGTTGTCGACATTCABAATGTCAATTTTTTTCGAATGAANAAT	2550
QY	2493	ACTTTCCCTTGTGTTTGGCAGGTTTGTAACTATTTATGAAGAAATATTTTAGCTGA	2552
DB	2551	ACTTTCCCTTGTGTTTGGCAGGTTTGTAACTATTTATGAAGAAATATTTTAGCTGA	2610
QY	2553	GTACTATATAATTTACAATCTTAAGAAATATCAAGTTGGGAACCAAGAAATAGCAAGG	2612
DB	2611	GTACTATATAATTTACAATCTTAAGAAATATCAAGTTGGGAACCAAGAAATAGCAAGG	2667
QY	2613	GAAATGTACAATTTTATCTTCTGGCAAGGACATCAATTCCTGTATTTATAGTGTATGTAA	2672
DB	2668	GAAATGTACAATTTTATCTTCTGGCAAGGACATCAATTCCTGTATTTATAGTGTATGTAA	2727
QY	2673	ATGCAACCTGTAAATGTTTACTTTTGTATTAATAATATGGGAGGGGGACTCAAAATTTCAGAAA	2732
DB	2728	ATGCAACCTGTAAATGTTTACTTTTGTATTAATAATATGGGA - GGGGAGCTCAAAATTTCAGAAA	2786
QY	2733	AGCTAAAAA	2741
DB	2787	AGCTACCAA	2795

O
 E
 F
 C
 D
 C

RESULTS
US-09-759-359A-3

US-09-739-359A-3
: Sequence 3, Application US/09759359A

: Sequence 3, Application of
: Patent No. US20020094560A1

;; PATENT NO. US2002009
: GENERAL INFORMATION:

APPLICANT: ABU-THREIDEH, Jane et al

APPLICANT: ABU-THREIDEH, Jame et al
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

1. TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NOVEL C

2. TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

RESULT 9

RESULTS
US-10-207-973-3

US-10-207-973-3
: Sequence 3. Application US/10207973

; Sequence 3, Application US/1020
; Publication No. US20030175927A1

```

; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/10/207,973
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 90541
; TYPE: DNA
; ORGANISM: Human
US-10-207-973-3

Query Match      20.5%; Score 666.2; DB 15; Length 90541;
Best Local Similarity 97.3%; Pred. No. 2.5e-147;
Matches 721; Conservative 0; Mismatches 13; Indels 7; Gaps 4;

QY 2004 AGGAGAACTGCACACATCACCACAGCTGAAGCCCTGGAGCCCTCTTTGATGTACTTGTGGA 2063
DB 87810 AGGAGAACTGCACACATCACCACAGCTGAAGCCCTGGAGCCCTCTTTGATGTACTTGTGGA 87869

QY 2064 AAGATATGCTGCGCCCATGAAGATGCTGCACAGTTTACAGATTTCCCTGATCCCGATGTT 2123
DB 87870 AAGATATGCTGCGCCCATGAAGATGCTGCACAGTTTCCCTGATCCCGATGTT 87929

QY 2124 AGAAATGTTTCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCCTTCGGCATCCTTGGTTGAA 2183
DB 87930 AGAAATGTTTCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCCTTCGGCATCCTTGGTTGAA 87989

QY 2184 TTCTTAGCAAAATCTACCAATATGCAATCTGAGCTAGCAAAATGTTCCAGTACATTGGA 2243
DB 87990 TTCTTAGCAAAATCTACCAATATGCAATCTGAGCTAGCAAAATGTTCCAGTACATTGGA 88049

QY 2244 CCTAAACGCTGACTCTCATCTTTTAAACAGGATTTACAAGTGAGCTGGCTTCCTCAGAC 2303
DB 88050 CCTAAACGCTGACTCTCATCTTTTAAACAGGATTTACAAGTGAGCTGGCTTCCTCAGAC 88109

QY 2421 ACAATAACATGCTCTTAAGCATTTCTTGTGTTGTGACATTTCAAAATGTCATTTTTT 2480
DB 88230 ACAATAACATGCTCTTAAGCATTTCTTGTGTTGTGACATTTCAAAATGTCATTTTTT 88289

QY 2481 TGAATGAAAAATACATTTCCCTTTGTTTTCGAGGTTTCTTAATAATTAAGGAA 2540
DB 88290 TGAATGAAAAATACATTTCCCTTTGTTTTCGAGGTTTCTTAATAATTAAGGAA 88349

QY 2541 TATTTTAGCTAGTACTATATAATTTACAATCTTAAGAAATTTACAAGTTGGAAACCAAG 2600
DB 88350 TATTTTAGCTAGTACTATATAATTTACAATCTTAAGAAATTTACAAGTTGGAAACCAAG 88409

QY 2601 AAAATAGCAAGGAAATGACAAATTTATCTTCTGCAAGGACATCATCTCTGTATTA 2660
DB 88410 AAAATAGCAAGGAAATGACAAATTTATCTTCTGCAAGGACATCATCTCTGTATTA 88466

QY 2661 TAGTGTATGTAATGCACCCTGTAATGTTACTTTGGATTAATATGGGAGGGGAGCTC 2720
DB 88467 TAGTGTATGTAATGCACCCTGTAATGTTACTTTGGATTAATATGGGAGGGGAGCTC 88525

QY 2721 AAATTTTCAGAAAAAGCTAAAAA 2741
DB 88526 AAATTTTCAGAAAAAGCTAAAAA 88546
```

```

RESULT 10
US-10-799-676-3
; Sequence 3, Application US/107999676
; Publication No. US20040157297A1
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001043DIV II
; CURRENT APPLICATION NUMBER: US/10/799,676
; CURRENT FILING DATE: 2004-03-15
; PRIOR APPLICATION NUMBER: 10/207,973
; PRIOR FILING DATE: 2002-07-31
; PRIOR APPLICATION NUMBER: 09/759,359
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 90541
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-799-676-3

Query Match      20.5%; Score 666.2; DB 17; Length 90541;
Best Local Similarity 97.3%; Pred. No. 2.5e-147;
Matches 721; Conservative 0; Mismatches 13; Indels 7; Gaps 4;

QY 2004 AGGAGAACTGCACACATCACCACAGCTGAAGCCCTGGAGCCCTCTTTGATGTACTTGTGGA 2063
DB 87810 AGGAGAACTGCACACATCACCACAGCTGAAGCCCTGGAGCCCTCTTTGATGTACTTGTGGA 87869

QY 2064 AAGATATGCTGCGCCCATGAAGATGCTGCACAGTTTACAGATTTCCCTGATCCCGATGTT 2123
DB 87870 AAGATATGCTGCGCCCATGAAGATGCTGCACAGTTTCCCTGATCCCGATGTT 87929

QY 2124 AGAAATGTTTCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCCTTCGGCATCCTTGGTTGAA 2183
DB 87930 AGAAATGTTTCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCCTTCGGCATCCTTGGTTGAA 87989

QY 2184 TTCTTAGCAAAATCTACCAATATGCAATCTGAGCTAGCAAAATGTTCCAGTACATTGGA 2243
DB 87990 TTCTTAGCAAAATCTACCAATATGCAATCTGAGCTAGCAAAATGTTCCAGTACATTGGA 88049

QY 2244 CCTAAACGCTGACTCTCATCTTTTAAACAGGATTTACAAGTGAGCTGGCTTCCTCAGAC 2303
DB 88050 CCTAAACGCTGACTCTCATCTTTTAAACAGGATTTACAAGTGAGCTGGCTTCCTCAGAC 88109

QY 2304 CTTTATTTTGTCTTTCAGGTAATGTTTGTGTTGTGACATTTTGTGCACTGTGATCCTG 2363
DB 88110 CTTTATTTTGTCTTTCAGGTAATGTTTGTGTTGTGACATTTTGTGCACTGTGATCCTG 88169

QY 2364 GGGAGGGGTAGTCTTTT--TGCTTTCAGCTAGTACTTTCACCATTTTCTT-CTGGAA 2420
DB 88170 GGGAGGGGTAGTCTTTTTCAGCTAGTACTTTCACCATTTTCTTCTTCTGGAA 88229

QY 2421 ACAATAACATGCTCTTAAGCATTTGTTTCTTGTGTTGTGACATTTCAAAATGTCATTTTTT 2480
DB 88230 ACAATAACATGCTCTTAAGCATTTGTTTCTTGTGTTGTGACATTTCAAAATGTCATTTTTT 88289

QY 2481 TGAATGAAAAATACATTTCCCTTTGTTTTCGAGGTTTCTTAATAATTAAGGAA 2540
DB 88290 TGAATGAAAAATACATTTCCCTTTGTTTTCGAGGTTTCTTAATAATTAAGGAA 88349

QY 2541 TATTTTAGCTAGTACTATATAATTTACAATCTTAAGAAATTTACAAGTTGGAAACCAAG 2600
DB 88350 TATTTTAGCTAGTACTATATAATTTACAATCTTAAGAAATTTACAAGTTGGAAACCAAG 88409

QY 2601 AAAATAGCAAGGAAATGACAAATTTATCTTCTGCAAGGACATCATCTCTGTATTA 2660
DB 88410 AAAATAGCAAGGAAATGACAAATTTATCTTCTGCAAGGACATCATCTCTGTATTA 88466

QY 2661 TAGTGTATGTAATGCACCCTGTAATGTTACTTTGGATTAATATGGGAGGGGAGCTC 2720
DB 88467 TAGTGTATGTAATGCACCCTGTAATGTTACTTTGGATTAATATGGGAGGGGAGCTC 88525

QY 2721 AAATTTTCAGAAAAAGCTAAAAA 2741
DB 88526 AAATTTTCAGAAAAAGCTAAAAA 88546
```

Db 88467 TAGTGTATGTAATGCACCTGTAAATGTTACTTTCCATTAAATATGGGA-GGGGACTC 88525
Qy 2721 AATTTTCAGAAAAGCTAAAAA 2741
Db 88526 AATTTTCAGAAAAGCTACCAA 88546

RESULT 11

US-09-880-107-3294
; Sequence 3294, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3294
; LENGTH: 4326
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 U09564
US-09-880-107-3294

Query Match 19.5%; Score 632.8; DB 9; Length 4326;
Best Local Similarity 60.4%; Pred. No. 3.5e-140;
Matches 1172; Conservative 0; Mismatches 672; Indels 96; Gaps 4;
Qy 250 CCGAGCCAGAGGAGGATCTGGGATCAGATGATGAGGCAAGAGGACCTCGGAC 309
Db 232 CCAGAGCAGGAAGAGGATCTGGGATCTGATGATGATGATGATGATGATGATGAT 291
Qy 310 TACTGCAAAAGGTGGATATCATCCAGTGAATAATGGAGACCTCTCAATGGCCGGTATCAT 369
Db 292 TATTGTAAAGGAGGTATCATCTTGTGAAAATTTGGAGATCTATTCAATGGGAGATACCAT 351
Qy 370 GTTATAGAAAGCTTGGATGGGGACCTTCTACTGTCTGCTGTGCTGGGATATGAG 429
Db 352 GTGATCCGAAAGTTAGGCTGGGGACACTTTTCAACAGTATGTTATCATGGGATATTCAG 411
Qy 430 GGGAAAAGATTTGTTGCAATGAAAGTTGTAAGAGTCCAGCATTTATACGGAGACAGCC 489
Db 412 GGGAGAAATTTGTGCAATGAAAGTAGTAAAGTGTGCAATTTACTGTAACAGCA 471
Qy 490 TTGGATGAAATAAAATTTGCTCAAAATGTTTCAGAAAAGTATCCAGTGAACCCAAACAAA 549
Db 472 CTAGATGAAATCCGGTTGCTGAAGTCAGTTGCAATTCAGACCTTAATGATCCAAATAGA 531
Qy 550 GACATGGTGGTCAGCTCATTCACGACTTCAGACTTCAGGATTCAGGATGATGATGATGATG 609
Db 532 GAAATGGTGTTCACACTACTAGTACTTTTAAATATCAGGAGTTAATGGAACACATATC 591
Qy 610 TGCATGGTCTTCAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAAAATCCAACTAT 669
Db 592 TGCATGGTATTTGAAGTTTGGGCAATCATCTGCTCAAGTGGATCATCAAAATCCAACTAT 651
Qy 670 CAAGGCCCTCCAGTACGTTGTGTAAGAGTATCATTCGACAGCTCTTCAAGGGTTAGAT 729
Db 652 CAGGGCTTCCACTGCTGTGTCAAAAAAATTTATTCAGCAAGTGTTCACAGGCTTTGAT 711
Qy 730 TACTTACAGTAAAGTGCAGATCATTCATCTACTGACATAAAGCCGAAAATATCTTGATG 789

Db 712 TATTATCATCAAGTCCCGTATATCCACTGACATTAATAAACAGAGAACATCTTATTG 771
Qy 790 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 849
Db 772 TCACTGAATGAGCAGTACATTTGGAGGCTGGTGCAGAAAGCAACAGAAATGGCAGCATCT 831
Qy 850 GGTGCT 909
Db 832 GGAGCT 891
Qy 910 AATATATCTAAAAACAAAAAGAAAAAATGAAAAAGAAAAACAGAAAGAGAGCGAGCTGAGTTA 969
Db 892 AAAATGTCAAAAGATAAAGAAAGAAATTTGAAGAAAGAGAGAGAGAGAGAGAGAGAG 951
Qy 970 TTGAGAGAGCGCTCTGAGGAGATAGAAATTTGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1029
Db 952 CTAGAGAGCGAATTCAGGAAATTTGAGAAATGAGAAAGAGAGTCTG----- 997
Qy 1030 GAAGAAAACATCACCTCAGCTGACCTTCCATGACAGAGTGGGATGAGTATGCTCCAGAG 1089
Db 998 -----GCCCTGGGCAAAAAAGACCAAAAC 1020
Qy 1090 GTGAAACTAAAAACAACAGAGATTAGAGAGCGCTGAGGCGAGAGATGCAAAAGGACAAAT 1149
Db 1021 AAGCAAGAGAAATCAGAGATCTCTGTTGAAAGACCTTTGAAGAGAGAGAGAGAGAGAG 1080
Qy 1150 GGTGAAGCTGAGGAT 1209
Db 1081 ATGACCCAGAGAAATTTGAAGAGTCAAGTACCATTTGGCCAGGATCAAACTTATGGAA 1140
Qy 1210 GAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1269
Db 1141 CGTGATACAGAGGCTGAGGAT 1200
Qy 1270 ACCAATGGCCATATTGAGATGGCCCTTCTCACTGGAGAGAGAGAGAGAGAGAGAGAGAT 1329
Db 1201 TATCTCGAAGACAGTATTAATGAACATTTGAGACATTAAGAGAGATCTACATAATGCTAAT 1260
Qy 1330 GATGATGAG 1389
Db 1261 GACTGTGATGTCGCAAAATTTGAATCAGGAATCTAGTTTCTTAACTCTCCCAATGGAGA- 1319
Qy 1390 AGTGATACACATATAGAGCTCTATGACAAATTCATGAGTGAATTTGCCAAATGGAGCA 1449
Db 1320 -----CAGCAGACATCTCAAGAAACAGACTCTTTGTACACCTATATAACA 1362
Qy 1450 CATAAATTCGCGAGTCACAGTTCGCGAGTTTCCACCTCGTTGTTCTCTGGATCCCTTA 1509
Db 1363 TCTGAGGTGTCAGACACATGTTGCGCAGTCTTCTCTCACTGTAGGTGAGTCAATTCAGT 1422
Qy 1510 GAACCTGTGGCTCGGCTCTGTGCTTTCTGAGGAGATCACCATTACTGAGAGAGAGAG 1569
Db 1423 GAAC-----AACATTTAGCCAACTTCAAGAAAGCAATTCGGGC 1460
Qy 1570 AGCAGTCCATCCATGACAG 1629
Db 1461 AGAGATACCTGTGAAAGATGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1515
Qy 1630 AAAGCAAAACCCCGGAGAGTGAATTTGTTGGTGAATCCCTGGATCCCGGGAATGCAGAT 1689
Db 1516 AAAGGAAATCCAGCGCTGGAAATTTCTTGTAACTCCCTTGGAGCAAAAGAGAGAGAG 1575
Qy 1690 AAAATTAGAGTAAATTTGCTGAGCTGGGAAATGCTTTGGTGGTGCATAAATCACTTCAG 1749
Db 1576 AAGCTCAAGGTGAAGATTTGCTGACCTTGGAAATGCTTTGGTGGTGCATAAATCACTT 1635
Qy 1750 GAAGACATCCAGAGCGCTCAGTACCGCTCCATAGAGGTTTTAAATAGAGAGCGGGTACAGC 1809
Db 1636 GAAGATATTCAAACAAAGGCAATATCGTTCTTTGGAAGTTCTAATCGGATCTGGCTATAT 1695
Qy 1810 ACCCTCGCGGACATCTGGAGCAGCGCTGTATGGCAATTTGAGCTTTGAGTGGCAACCGGAGAT 1869
Db 1696 ACCCTCGTGCATTTGGAGCAGCGCATGCTGGCTTTTGAACCTGGCCACAGGTGACTAT 1755

QY 1870 TTGTTTGAACCACTTCTGGGAAGACTATTCCAGAGCGAAGACCACATAGCCACATC 1929
|||||
DB 1756 TTGTTTGAACCTCATTCAGGGAAGAGTACACTCGAGATGAAGATCACATTCATTTGATC 1815
|||||
QY 1930 ATAGAGCTGCTAGGCAAGTATTTCCAAAGGCACTTTTGTCTCTATCTGAAATATTTCTCGGAA 1989
|||||
DB 1816 ATAGAACTTCTGGGAAGGTGCTCGCAAGCTCATTTGTGGCAGGAAATATTTCCAGGAA 1875
|||||
QY 1990 TTCTTCAATCGCAGAGAGAACTGCGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTTT 2049
|||||
DB 1876 TTTTTCACCAAAAAGGTGACCTGAAACATATCACGAAGCTGAAACCTTTGGGCTTTT 1935
|||||
QY 2050 GATGTACTTGTGAAAGTATGCTGCGCCCATGAAGATGCTGCACAGTTTACAGATTTC 2109
|||||
DB 1936 GAGGTTCTAGTGAGAAGTATGATGGTCTCAGGAAGAGCAGCTGGCTTCACAGATTTC 1995
|||||
QY 2110 CTGATCCGATGTTAGAAATGGTTCCAGAAACAGAGCCTCAGCTGGCGAATGCCCTTCGG 2169
|||||
DB 1996 TTACTGCCCATGTTGGAGCTGATCCCTGAGAGAGAGCCACTGCGCGGAGTGTCTCCGG 2055
|||||
QY 2170 CATCCTTGGTTGAATCTTA 2189
|||||
DB 2056 CACCCCTTGGCTTAACCTTA 2075
|||||

RESULT 12

US-10-172-118-875
; Sequence 875, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yundong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 875
; LENGTH: 4326
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM 003137
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-875

Query Match 19.5%; Score 632.8; DB 15; Length 4326;
Best Local Similarity 60.4%; Pred. No. 3.5e-140;
Matches 1172; Conservative 0; Mismatches 672; Indels 96; Gaps 4;
QY 250 CCGAGCCAGAGGAGAGATCTGGGATCAGATGATGAGGCAAGAGACCCTGGCGAC 309
|||||
DB 232 CNAGACGAGAGAGAGATCTGGGATCTGATGATGAGCAGAGATCTTAATGAT 291
|||||
QY 310 TACTGCAAGGTGGATPATCATCCAGTGAAATTTGGAGACCTCTTCAATGGCGGTATCAT 369
|||||
DB 292 TATTGTAAGGAGTTATCATCTTGTGAAATTTGGAGATCTATTCAATGGAGATACCAT 351
|||||
QY 370 GTTATTAGAAAGCTTGATGGGGCACTTCTCTACTGTCTGGCTGTGGGATATGACAG 429
|||||
DB 352 GTGATCCGAAAGTTAGGCTGGGACACTTTTCAACAGATATGTTATCATGGGATATTCAG 411
|||||
QY 430 GCGAAAGATTTCTTCAATGAAGTTGTAAGAGTCCCGAGCATTTATACGAGACAGCC 489
|||||

DB 412 GCGAAGAAATTTGTGCAATGAAAGTAGTTAAAGTGTGAACATTTACACTGAAACAGCA 471
QY 490 TTGGAATGAAATAAAATTTGCTCAATGTGTTCCAGAAAGTGTATCCAGTAGTCCCAACAA 549
|||||
DB 472 CTAGATGAAATCCGGTTGCTGAAGTCAGTTCCGAATTTAGAGCCCTAATGATCCAAATAGA 531
|||||
QY 550 GACATGGTGGTCCAGCTCAATTCAGCACTTCAAGATTTTCAAGGATGAATGGATACATGTC 609
|||||
DB 532 GAAATGGTGTTCAACTACTAGATGACTTTAAATATCAGGAGTTAATGGAACACATATC 591
|||||
QY 610 TGCATGGTCTTCCGAAGTACTTGGCCCACTCTCTCAAGTGGATCATCAAAATCCAACTAT 669
|||||
DB 592 TGCATGGTATTTTGAAGTTTGGGCACTCATCTGCTCAAGTGGATCATCAAAATCCAATAT 651
|||||
QY 670 CAAAGGCTCCAGTAGTGTGTGAAGATATCATTCGACAGGTCTCTCAAGGGTTAGAT 729
|||||
DB 652 CAGGGGCTTCCACTGCTTGTCTCAAAAAAATTTATTCAGCAAGTGTGTA CAGGGTCTTGAT 711
|||||
QY 730 TACTTTACACAGTAAGTGAAGATCATTTATCTGACATAAAGCCGAAATATCTTTGATG 789
|||||
DB 712 TATTTACATACCAAGTGGCGTATCATCCACTGACATTTAAACAGAGAACATCTTTATG 771
|||||
QY 790 TGTGTGGATGATGCATATGTGAGAAGAAATGGCAGCTGAGGCCCACTGAGTGGCAGAAAGCA 849
|||||
DB 772 TCAGTGAATGACAGGTACATTCGAGGCTGGCTGCAGAAGCAACAGAAATGGCAGCGATCT 831
|||||
QY 850 GGTGCTCTCTCTCTTCCAGGTCTGCGAGTAGTACGGCTCCACAGCAGAAACCTATAGGA 909
|||||
DB 832 GGAGCTCTCTCCGCTTCCGGATCTGCAGTCAGTACTGCTCCCGAGCTTAAACAGCTGAC 891
|||||
QY 910 AAAATATCTAAAAAAGAAAGAACTGAAAGAAAGAAACAGAAAGAGAGGAGCTGAGTTA 969
|||||
DB 892 AAAATGTCAAGAAATGAAGAAAGAAATTTGAAGAAAGACAGAGCCCGGCGAGAAATTA 951
|||||
QY 970 TTGAGAGAGCGCTGCGAGGAGATAGAAGAAATTTGGAGCGAGAGCTGAAAGGAAATAA 1029
|||||
DB 952 CTAGAGAGCGAATGCGAGAAATTTGAGAAATTTGAGAAAGAGTCCG----- 997
|||||
QY 1030 GAAGAAACATCACCTCAGCTGACCTTCCAATGACCAGATGGCGAATATCTGCCAGAG 1089
|||||
DB 998 -----GCCCTGGGCAAAAGACCAAAAC 1020
|||||
QY 1090 GTGAAACTAAAAACAACAGGATTTAGAGGAGCGGCTGAGCGAGAGACTGCAAAAGGACAAT 1149
|||||
DB 1021 AAGCAAGAAAGATCAGAGAGTCTGTTGAAGACCCTTGAAGAGAGAACCCACCTAATAAA 1080
|||||
QY 1150 GGTGAAGCTGAGGACCAGGAGAGAAAGATGCTGAGAAAGAAACATTTGAAAGAGAT 1209
|||||
DB 1081 ATGACCCAGAAAAACTTGAAGAGTCAAGTACCATTGGCCAGGATCAAAACGCTTATGGAA 1140
|||||
QY 1210 GAAGATGATGTAGATCAGGAACTTGGGAACATAGACCCTACGTGGATAGAAATCACCTAAA 1269
|||||
DB 1141 CGTGATACAGAGGGTGGTGCAGCAGAAATTAATTTGCAATGGAGTGTGTAAGTCAATTAAT 1200
|||||
QY 1270 ACCAATGGCCATTTTGAAGATGGCCATTTCTCACTGGAGCAGCACTGGACGATGAAGAT 1329
|||||
DB 1201 TATACTCAGAACAGTAATAATAAACAATTGAGACATPAAGAGGATCTACATAATGCTAAT 1260
|||||
QY 1330 GATGATGAAGAAAGACTGCCCCAAATCCTGAGGAATAATAATCTTGTATGAGCCAAATGCAGAA 1389
|||||
DB 1261 GACTGTGATGTCCAAAATTTGAATCAGGAATCTAGTTTCTTAAGTCTCCCAATGGAGA- 1319
|||||
QY 1390 AGTGATTAACATA TAGCAGCTCTTATGAACAAATTTCAATGGTGAATTTGCCAAATGGACGA 1449
|||||
DB 1320 -----CAGCAGACATCTCAAGAAACAGACTCTTTGTACACCTATAAACA 1362
|||||
QY 1450 CATAAATTTCCGAGTCAAGTTCCAGAGTTTCCACCTCGTTGTTCTCTGATCCTTA 1509
|||||
DB 1363 TCTGAGGTTGACAGACCACTAGTGTGCCAGTCTTCTCAACTGTAGGTGATCATTCAGT 1422
|||||
QY 1510 GAACTGTGGCTGCGGCTCTGTGCTTTCTGAGGGATCAACCACTTACTGACGAAGAGGAG 1569
|||||
DB 1423 GAAC-----AACACATTACGCAACTTCAAGAAAGCATTCGGGC 1460
|||||

		; ORGANISM: Homo sapiens	
		US-10-342-887-875	
		Query Match 19.5%; Score 632.8; DB 16; Length 4326;	
		Best Local Similarity 60.4%; Pred. No. 3.5e-140; Indels 96; Gaps 4;	
		Matches 1172; Conservative 0; Mismatches 672;	
QY	1570	AGCAGTCCATCCATGACAGAGACGAAACGGTTTCAGCTCAGTACTCGGGATTGGCA	1629
DB	1461	AGAGATACCTGTGAAGATGAACAGAGCAAGAAATAAACGACCACTGGACA-----AC	1515
QY	1630	AAAGCAAAACCCGGGAGCTGACTTGTGTGTAATCCCTGGATCCCGGAAATGCAGAT	1689
DB	1516	AAAGGAAATCCACGGCTGGAAATTTCTTGTTAATCCCTTGAGCCAAATAATGCAGAA	1575
QY	1690	AAAAATTAGAGTAAAAATTGCTGACCTCGGAAATGCTTGTGGGTGCATAAACTTCACG	1749
DB	1576	AGCTCAAGTGAAGATTGCTGACCTTGGAATGCTTGTGGTGCACAAACATTTCACT	1635
QY	1750	GAAGACATCCAGCGCTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGC	1809
DB	1636	GAAGATATTCAAAACAGGCAATATCGTTCTTGGAAAGTTCTAATCGGATCTGGCTATAAT	1695
QY	1810	ACCCCTGGGACATCTGGAGCACGGCTGTATGCAATTTGAGCTGGCAACGGGAGATTAT	1869
DB	1696	ACCCCTGTGACATTTGGAGCACGGCATGCTATGCTTTGAACTGGCCACAGGTGACTAT	1755
QY	1870	TTGTTTGAACCACTTCTCGGGAAGACTATTCCAGAGACGAAGACCAATAGCCCAATC	1929
DB	1756	TTGTTTGAACCTCAITTCAGGGGAAGATACACTCGAGATGAAGATCACATTGCAATTGATC	1815
QY	1930	ATAGAGCTGTAGGAGTATTCAGGCACTTTGCTCTATCTGGAAATATTTCTCGGAA	1989
DB	1816	ATAGAACTTCTGGGGAAGTGCTCGCAAGCTCAITTTGGGAGGAAATATTTCAAAGAA	1875
QY	1990	TTCTTCAATCGCAGAGAGAACTGCGACACATCACCAAGCTGAAGCCCTGGAGCCCTTTT	2049
DB	1876	TTTTTCAACAAAAGGTGACCTGAAACATAATACCAAGCTGAACCTTTGGGGCTTTT	1935
QY	2050	GATGTACTTGTGAAAGATATGCTGGCCCATGAAGATGCTGCAAGTTTACAGATTTC	2109
DB	1936	GAGGTTCTAGTGAGAAGTATGAGTGGTCTCAGGAGAGAGCGAGCTGGCTTTCAGATTTC	1995
QY	2110	CTGATCCGATGTTAGAAATGGTTCCAGAAAAACGAGCTCAGCTGGCGAATGCCCTTCGG	2169
DB	1996	TTACTGCCCATGTTGAGCTGATCCCTGAGAGAGAGCCACTGCGCCGAGTGTCTCCGG	2055
QY	2170	CATCCTTGGTTGAATTTCTTA	2189
DB	2056	CACCTTGGCTTAATCTCTA	2075
RESULT 13			
US-10-342-887-875			
; Sequence 875, Application US/10342887			
; Publication No. US20040058340A1			
; GENERAL INFORMATION:			
; APPLICANT: Dai, Hongyue			
; APPLICANT: He, Yudong			
; APPLICANT: Linsley, Peter S.			
; APPLICANT: Mao, Mao			
; APPLICANT: Roberts, Christopher J.			
; APPLICANT: Van 't Veer, Laura Johanna			
; APPLICANT: Van de Vijver, Marc J.			
; APPLICANT: Bernards, Rene			
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients			
; FILE REFERENCE: 9301-188-999			
; CURRENT APPLICATION NUMBER: US/10/342.887			
; CURRENT FILING DATE: 2003-01-15			
; PRIOR APPLICATION NUMBER: 60/298,918			
; PRIOR FILING DATE: 2001-06-18			
; PRIOR APPLICATION NUMBER: 60/380,710			
; PRIOR FILING DATE: 2002-05-14			
; PRIOR APPLICATION NUMBER: 10/172,118			
; PRIOR FILING DATE: 2002-06-14			
; NUMBER OF SEQ ID NOS: 2699			
; SEQ ID NO 875			
; LENGTH: 4326			
; TYPE: DNA			

Db 1141 CGGTATACAGAGGGTGGTGCAGCAGAAATTAATTGCAATGGAGTGATTGAAGTCATTAAT 1200
Qy 1270 ACCAATGGCCATATTAGAAATGSCCACTTCTACTGGAGCAGCAACATGACGATGAAGAT 1329
Db 1201 TATACTCAGAACAGTAATAATGAACATTAAGAGAGGATCTACATATGCTAAT 1260
Qy 1330 GATGATGAAGAGAGCTGCCAAATCTGAGGAATAAATCTTGATGAGCAAAATGAGAA 1389
Db 1261 GACTGTGATGTCAAAATTTGAATCAGGAATCTAGTTTCTTAAGTCTCCAAATGGAGA- 1319
Qy 1390 AGTGATTACACATPATAGCAGCTCTATGAACAAATCAATGGTGAATGGCAATGACGA 1449
Db 1320 -----CAGCAGCACATCTCAAGAAACAGACTCTTGACACCTATAACA 1362
Qy 1450 CATAAATTTCCCGAGTACAGTTCCAGAGATTTTCCACCTCGTTGTCTCTGATCCTTA 1509
Db 1363 TCTGAGGTGTCAGACACCATGTTGTCGCACTCTTCTCACTAGTCTAGTCTAGTCAATCAGT 1422
Qy 1510 GAACCTGTGGCCTGCGGCTCTGTGCTTCTGAGGATCACCACCTTACTGAGCAAGAGGAG 1569
Db 1423 GAAC-----AACATTAGCCAACTTCAAGAAAGCATTCGGGC 1460
Qy 1570 AGCAGTCCATCCCATGACAGAGCAGACGGTTTTCAGCCTCCAGTACTGGGATTTGCCA 1629
Db 1461 AGAGATACCTGTGAAGATGNAACAAGCAAGAACATAACGGACCACTGGACA-----AC 1515
Qy 1630 AAGCAAAAACCGGCGAGCTGACTTGTGTTGTAATCCCTCGATCCGCGGAATGAGAT 1689
Db 1516 AAGGAAAAATCCACGGCTGGAATTTTCTTGTAAATCCCTTTGAGCCAAAATGAGAA 1575
Qy 1690 AAAATTTAGATGTAATAATCTGACCTGGGAATGCTTGTGGTGATAAACACTTCAGC 1749
Db 1576 AAGCTCAAGGTGAAGATGCTGACCTTGGAAATGCTTGTGGTGACAAACATTTCACT 1635
Qy 1750 GAAGACATCCAGACGGCTCAGTACCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGC 1809
Db 1636 GAAGATATTCAAACAGGCAATATCGTCTTGGAGTTCTAATCGATCTGGCTAAT 1695
Qy 1810 ACCCTGCGGACATCTGGAGCAGCGGTGTATGGCAATTTGAGCTGGCAACGGGAGATTA 1869
Db 1696 ACCCTGCTGACATTTGGAGCAGCGCATGCTGGCCCTTTGAACTGGCCACAGGTGACTAT 1755
Qy 1870 TTGTTTGAACCATCTCGGGAAGACTATTCAGAGAGCAGACACACATAGCCACATC 1929
Db 1756 TTGTTTGAACCTCATTCAGGGAAGAGTACCTCGAGATGAAGATCACATGCTATGATC 1815
Qy 1930 ATAGAGCTGTAGGCAAGTATTCGAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGAA 1989
Db 1816 ATAGAATCTTCTGGGAGGTGCTCCAGCTCATTTGTGGCAGGAAATATTCGAAGAA 1875
Qy 1990 TTCTTCAATCGCAGAGAGAACTGGACACATCACCAAGCTGAAGCCCTGGAGCCTCTTT 2049
Db 1876 TTTTTCACCAAAAAGGTGACCTGAAACATATCACGAAGCTGAACCTTTGGGCGCTTTT 1935
Qy 2050 GATGTACTTGTGAAAAGTATGCTGGCCCAATGAAGATGCTGCACAGTTTACAGATTC 2109
Db 1936 GAGGTTCATGTGAGAAATGATGATGCTGCTCAGGAAGAGGCACTGGCTTTCAGATTT 1995
Qy 2110 CTGATCCCATGTTAGAAATGGTTCCAGAAAAACGAGCCTCAGCTGGCGAATGCCCTCGG 2169
Db 1996 TTACTGCCCATGTTGAGCTGATCCCTGAGAGAGGACCACTGCGCCGAGTGTCTCCGG 2055
Qy 2170 CATCTTGGTTGAATTTCTTA 2189
Db 2056 CACCCCTTGGCTTAACCTCTA 2075

RESULT 14

US-10-252-157-374
; Sequence 374, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary

; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 374
; LENGTH: 4349
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 474724.5
US-10-252-157-374

Query Match 19.4%; Score 629.6; DB 15; Length 4349;
Best Local Similarity 60.3%; Pred. No. 2.1e-139;
Matches 1170; Conservative 0; Mismatches 674; Indels 96; Gaps 4;

Qy 250 CCGGAGCCAGAGAGAGAGATCCTGGGATCAGATGATGAGGAGCAAGAGGACCCCTGGGAC 309
Db 232 CCAGAGCAGGAAGAGAGATCTGGGATCTGATGATGATGAGCAAGAGATCCTAATGAT 291
Qy 310 TACTGCAAAAGGTGGATATCATCCAGTGAATAATGGAGACCTCTTCAATGGCCGGTATCAT 369
Db 292 TATTGTAAAGAGAGGTATCATCTTGTGAAAAATGGAGATCTATTCAATGGAGATACCAT 351
Qy 370 GTTATTAGAAAGCTTGGATGGGGGACATTTCTACTGTCTGCTGTGCTGGGATATGCG 429
Db 352 GTGATCCGAAGGTAGGCTGGGGACATTTTCAACAGATATGTTATCATGGGATATTCAG 411
Qy 430 CGGAAAAAGATTTTGTGCAATGAAAGTTGTAAAAAGTGCACGATTTATACGGAGACAGCC 489
Db 412 GGGAGAAATTTTGTGGCAATGAAAGTAGTTAAAGTCTGAACATTTACACTGAAACAGCA 471
Qy 490 TTGGATGAATAAATTTGCTCAAAATGTGTTTCGAGAAAGTGATCCCAAGTACCCAAACAAA 549
Db 472 CTAGATGAATCCGGTTGCTGAAGTCAGTTCCGAATTCAGACCCATTAATGATCCAAATAGA 531
Qy 550 GACATGGTGGTCCAGCTCATTCAGCAGCTTCAAGATTTTCAGGCATGATGGGATACATGTC 609
Db 532 GAAATGGTGTGTTCAACTACTAGTACATTTTAAATATCAGGAGTTAATGGACACATATC 591
Qy 610 TGCAATGGTCTTCCAGATGATCTGGCCACCATCTCTCAAGTGGATCATCAATCAACTAT 669
Db 592 TGCATGGTATTTGAAGTTTGGGGCATCATCTGCTCAAGTGGATCATCAATCCATTTAT 651
Qy 670 CAAGGCTCCCAAGTACGTTGTGGAAGAGTATCATTCGACAGGTCTCTCAAGGGTTAGAT 729
Db 652 CAGGGCTTCCACTGCTTGTGTCAAAAAAATTTATTCAGCAAGTGTTCACAGGCTCTGAT 711
Qy 730 TACTTACACAGTAAGTGCAGATCATTTCACTCATGACATTAAGCCGGAATAATCTTGATG 789
Db 712 TATTTACATACCAAGTGGCGGTATCATCCACATCTGACATTTAAACACAGAGAACATCTTAT 771
Qy 790 TGTGTGGATGATGATATGTTGAGAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCA 849
Db 772 TCAGTGAATGACAGTATCATTCGGAGGCTGGCTGCAGAGCAACAGATGGCAGCGATCT 831
Qy 850 GGTGCTCTCTCTCTCTTCAAGGCTGTCAGTGAAGTACGGCTCCACAGAGAAACCTATAGGA 909
Db 832 GGAGCTCTCTCGCTTCCGGATCTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 891
Qy 910 AAAATATCTAAAAACAAAGAAAAAAGTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAGAA 969
Db 892 AAAATGTCAAGAAATAAGAAAGAAATTTGAAGAAAGTGAAGAAAGTGAAGAAAGTGAAG 951
Qy 970 TTGGAGAAAGCGCTGTCAGGAGATAGAAAGATTTGGAGCGAGAGAGTCTCAAGAGAAATA 1029
Db 952 CTAGAGAACGGAATCAGGAAATTTGAGGAAATTTGAGGAAATTTGAGGAAATTTGAGGAA 997

Qy 1030 GAAGAAAACATCACTCAGCTGCACCTTCCATGACACGAGTGGCGAATATCTGCCAG 1089
Db 998 -----GCCCTGGCGCAAAAAGAACCAAC 1020
Qy 1090 GTGAAGAACTAAACCAACAGGATTTAGAGGAGCGGCTGAGGCGAGAGACTGCAAGAGCAAT 1149
Db 1021 AAGCAGAAGAAATCAGAGAGTCTCTTGAAGACCTTTGAAAGAGACCCACCTAATAA 1080
Qy 1150 GGTGAAGCTGAGGACGAGAGAGAAAGAGATGCTGAGAAAGAAAAATTTGAAAGAT 1209
Db 1081 ATGACCCAAAGAAAACCTTGAAGAGTCAAGTACCATTTGGCCAGGATCAAAACGCTTATGGAA 1140
Qy 1210 GAAGATGATGATAGATCAGAACTTGGACATAGACCTTACGTGATAGATCACTAAA 1269
Db 1141 CGTGATACAGAGGCTGTCAGCAGAAATTAATGCAATGGAGTATTTGAAGTCAATTAAT 1200
Qy 1270 ACCAATGGCCATATGAGAAATGCCCATTTCTCACTGGAGCAGCAACTTGACCAATGAAGAT 1329
Db 1201 TATACTCAGACGATTAATGAACATTTGACATAGAGAGATCTACATAATGCTAAT 1260
Qy 1330 GATGATGAAGAGACTGCCCAATCTCTGAGGAATATAATCTTGATGAGCCAAATGCAGAA 1389
Db 1261 GACTGTGATGTCAAAATTTGAATCAGGAATCTAGTTTCTTAAGCTCCCAAAATGGAG- 1319
Qy 1390 AGTGATACATATAGCAGCTCTATGAACAATTCATGGTGAATTTGGCAATGGACGA 1449
Db 1320 -----CHAGCAGCATCTCAAGAAACAGACTTGTGTACACTATNACA 1362
Qy 1450 CATAAAATCCCGAGTACAGATTTCCAGAGATTTCCACCTGTTGTTCTCTGGATCCTTA 1509
Db 1363 TCTGAGGTGTCAGACACCATGGTGTGCGATCTTCTCACTGTAGGTCACTCAATTCAGT 1422
Qy 1510 GAACCTGTGGCTGGCTGTGCTTTCTGAGGATCACCATTACTGAGCAAGAGGAG 1569
Db 1423 GAAC-----AACACATTAGCCAACTTCAAGAAAGCATTCGGGC 1460
Qy 1570 AGCAGTCCATCCATGACAGAGACGATTTCCAGCTTCCAGTACTGGGATTTGCCA 1629
Db 1461 AGATATACCTGTGAGATGACACAGAGCAAGACATAACGGACCACTGGACA-----AC 1515
Qy 1630 AAAGCAAAAACCCGGGAGCTGATCTTGTGGTGAATCCCTCGATCGCGGATGAGAT 1689
Db 1516 AAAGAAAATCCACGGCTGGAATTTTCTTGTGTAATCCCTTGGCCCAAAAATCAGAA 1575
Qy 1690 AAAATTAGATGAATTTGTCACCTGGGAAATGCTTGTGGTGCATAAACACTTCAAG 1749
Db 1576 AAGCTCAAGGTGAAGATTGCTGACCTTGGAAATGCTTGTGGTGCACAAACATTTCACT 1635
Qy 1750 GAAGACATCCAGACGCTCAGTACCGCTCCATAGAGTTTAAATAGAGCGGGGTACAGC 1809
Db 1636 GAAGATATCAACAAGGCAATATCGTTCTTGGAGTTCTTAATCGGATCTGGCTATAAT 1695
Qy 1810 ACCCTGGGACATCTGAGACAGCGCTGTATGCTATTCGATTCGCTGGCAACGGGAGATTAT 1869
Db 1696 ACCCTGTGTGATTTGAGCAGCGCATGATGCGCTTTGAACCTGGCCACAGGTGACTAT 1755
Qy 1870 TTGTTGAACACATTTCTGGGGAAGACTATTCCAGACAGCAAGACCATAGCCCAATC 1929
Db 1756 TTGTTGAACTCATTCAGGGAAGAGTACACTCGAGATGAAGATCACTGATGATGATC 1815
Qy 1930 ATAGAGCTGTAGGAGATTTCCAGGCACTTCTCTATCTGGAATATTTCTCGGGAA 1989
Db 1816 ATAGAACTTCTGGGGAAGTGCCTCGCAAGCTCATTTGTCAGGAGAAAATATTCCAAGGA 1875
Qy 1990 TCTTCAATCCGAGGAGACTCGACATATCAACAGCTGGAAGCTTGGAGCTCTCTT 2049
Db 1876 TTTTTCACAAAAGGTGACCTGAAACATATACAGAGCTGAAACCTTGGGCTTTT 1935
Qy 2050 GATGTACTTGTGAAAAGTATGGCTGGCCCATGAAGATGCTGCACAGTTTACAGATTTC 2109
Db 1936 GAGGTTCTAGTGGAGAGTATGATGCTCGCAGGAAGAGGAGCTGGCTTACAGATTTC 1995

Qy 2110 CTGATCCCGATGTAGAAATGGTTCCAGAAAACGAGCCCTCAGCTGGCGAATGCCCTTGG 2169
Db 1996 TTAAGTCCCATGTTGGAGCTGATCCCTGAGAGAGAGCCACTGCCCGCGAGTGTCTCCGG 2055
Qy 2170 CATCCTTGGTTGAATCTTA 2189
Db 2056 CACCTTGGCTTAACCTCTA 2075
RESULT 15
US-09-814-353-21987
; Sequence 21987, Application US/09814353
; Publication No. US20030165831A1
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21987
; LENGTH: 4639
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 4427, 4428, 4429, 4430, 4431, 4432, 4433, 4434, 4435, 4436,
; LOCATION: 4437, 4438, 4439, 4440, 4441, 4442, 4443, 4444, 4445, 4446,
; LOCATION: 4447, 4448, 4449, 4450, 4451, 4452, 4453, 4454, 4455, 4456,
; LOCATION: 4457
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-21987
Query Match 19.4%; Score 629.6; DB 10; Length 4639;
Best Local Similarity 60.3%; Pred. No. 2.2e-139;
Matches 1170; Conservative 0; Mismatches 674; Indels 96; Gaps 4;
Qy 250 CCGAGAGCCAGAGAGAGATCCTGGGATCAGATGATGAGAGCAAGAGACCTCGGAC 309
Db 220 CAGAGCAGAGAGAGAGATTCCTGGATCTGATGATGATGAGCAAGAGATCCTAATGAT 279
Qy 310 TACTGCAAGGTGGATATCATCCAGTGAATTTGAGACCTTTCATGCGCGGTATCAT 369
Db 280 TATTGTAAGAGGTTATCATCTTGTGAAATTTGAGATCTATTCAATGGGAGATCACCAT 339
Qy 370 GTTATTAGAAAGCTTGGATGGGGCACTTCTACTGTCTGCTGTGCTGGGATATGAG 429
Db 340 GTGATCCGAAAGTTAGCTGGGACACTTTTCAACAGTATGTTATCATGGGATATTCAG 399
Qy 430 GGGAAAGATTTTGTGCAATGAAGTTGFAAAGTGTCCAGCATATTACGAGAGACGC 489
Db 400 GGGAGAAATTTGTGGCAATGAAGTAGTTAAAGTAGTGAACATTTACACTGAACAGCA 459
Qy 490 TTGATGAATAAATTTGCTCAATGCTTTCAGAGAAAGTATCCAGTGCACCAACAAA 549
Db 460 CTAGATGAATCCGGTTGCTGAAGTCAATTCGCAATTCAGACCTTAATGATCCNAATAGA 519

Qy	550	GACATGGTGTGCTCAATTGACGACTTTCAAGATTTTCAGGCATGAATGGGATACATGTC	609
Db	520	GAATGGTGTGTTCAACTACTAGATGACTTTAAAAATATCAGGAGTTAATGGAAACAATATC	579
Qy	610	TGCATGCTTTCGAAGTACTTGGCCACCATCTCTCTCAAGTGGAATCATCAAAATCCAATAT	669
Db	580	TGCATGGTATTTGAAGTTTTGGGCGATCACTGCTCAAGTGGAATCATCAAAATCCAATAT	639
Qy	670	CAAGGCCCTCCCAGTAGTACGTTGTGTGAAGAGATCAATTCGACAGGTCCCTTCAAGGGTTAGAT	729
Db	640	CAGGGCTTCCACTGCCTTGTGTCAAAAATTAATTCAGCAAGTGTTCACGGGTCTTGAT	699
Qy	730	TACTTACACAGTAGTGCAAGATCAATTCATATCTGACATATAAGCCGGAAAAATATCTTGAATG	789
Db	700	TATTTACATACCAAGTSCCGTATCATCCACATGACATTAACCCAGAGAACAATCTTTATTG	759
Qy	790	TGTGTGATGATGCATATGTGTGAGAAGAAATGCGACCTGAGGCCACTGAGTGGCAGAAAGCA	849
Db	760	TCAGTGAATGAGCAGTACATCTCGAGAGCTGGCTGCAGAGCAACAGAAATGGCAGCGATCT	819
Qy	850	GGTGCTCCTCTCTCTCAGGGTCTGCAGTGAGTAGCGGCTCCACAGCAGAAACCTATAGGA	909
Db	820	GGAGCTCTCCGCCCTTCGGGATCTGCAGTCACTACTGCTCCCCAGCCTAAACAGAGTAC	879
Qy	910	AAAATATCTAAAACAAAAGAAAATACTGAAAAGAAAACAGAAGAGCGCTGAGTTA	969
Db	880	AAAATGTCAAGAATAAAGAAAGAAAATGAAGAAGAACGACAGACGCCAGGCAGCAATTA	939
Qy	970	TTGAGAGAAGCGCTTCAGGAGATAGAAGAAATTTGAGCGAGAAGCTGAAAAGSMAAATAA	1029
Db	940	CTAGAGAAGCGAATGAGGAAATTTGAGAAATGAGAAAGNGTCGG-----	985
Qy	1030	GAAGAAACATCACCTCAGCTGCACCTTCCAAATGACCAGGATGCGAATATCTGCCACAG	1089
Db	986	-----GCCCTGGGCAAAAAGACCAAAAC	1008
Qy	1090	GTGAAACTPAAAAACAAACAGANTTAGAGAGGCGGCTGAGGCAGNAGACTGCAAGGACAAT	1149
Db	1009	AAGCAAGAGAATCAGAGAGTCTGTTGAAAGACCTTTGAAAGAGAAACCCACCTTAATAA	1068
Qy	1150	GGTGAAGCTGAGGACAGGAGAGAAAGAAAGATGCTGAGAAAGAAAAACATTTGAAAAAGAT	1209
Db	1069	ATGACCCAAAGAAAACCTTGAAGAGTCAAGTACCATTTGCCAGGATCAACGCTTATGAA	1128
Qy	1210	GAGATGATGTAGATCAGGAACTTTGGAAACATAGACCCCTGCTGGATAGAAATCACCTTAA	1269
Db	1129	CGTGATACAGAGGGTGTGCAGCAGAAATTAATTTGCAATGGAGTGATTGAAGTCATTAAT	1188
Qy	1270	ACCAATGGCCATATTGAGATGGCCCAATCTCACTGGAGCAGCAACTGGACGATGAAGAT	1329
Db	1189	TATACTCAGAACAGTAATAATGAACANITGAGACATAAAGAGGATCTACATAATGCTAAT	1248
Qy	1330	GATGATGAAGAAGACTCCCAAAATCCTGAGGAATATAATCTTGATGAGCCAAATGCGAA	1389
Db	1249	GACTGTGATGTCCAAAATTTGAATCAGGAATCTAGTTTCTTAAGTCCCAAAATGGAGA-	1307
Qy	1390	AGTGATTACATATAGCAGCTCTTATGAAACAAATTAATGGTGAATTTGCCAAATGGACGA	1449
Db	1308	-----CAGCAGCACATCTCAAGAAAACAGACTCTTGACACCTATATAACA	1350
Qy	1450	CATAAAATTCGGAGTCACAGTTCACAGATTTTCCACCTCGTTGTTCTCTGGATCTCTTA	1509
Db	1351	TCTGAGTGTGAGACACATGGTGTGCCAGTCTTCTCCTCAACTGTGATGATGATCAGT	1410
Qy	1510	GAACCTGTGGCCTCGGCTCTGTGCTTTCTGAGGGATCACCACTTACTGACGAGAGGAG	1569
Db	1411	GAAC-----AACACATTTAGCCAACTTCAAGAAAGCATTTCCGGC	1448
Qy	1570	AGCAGTCCATCCATGACAGAACAGAAACGGTTTTCAGCTCCAGTACTGGGATTTGCCCA	1629
Db	1449	AGAGATACCTGTGAAGATGAAACAGAGCAGAAACATAACGGACCACTGGACA-----AC	1503

QY	1630	AAAGCAAAAACCCGGGAGCTGACTTGTTCGGTGAATCCCTCTGGATCCGCGGAATGCAGAT	1689
DB	1504	AAAGGAAAAATCCACGGCTGGAAATTTTCTTGTTAATCCCTTGAGCCAAAAAATTCGAGAA	1563
QY	1690	AAAATTAGAGTAAAAATTTGCTGCACCTCGGAAATGCTTGTGGGTGCAATAAACACTTCCAG	1749
DB	1564	AAGCTCAAGGTGAAGATTGCTGCACCTTTGGAATATGCTTGTGGGTGCAACAACATTTCACT	1623
QY	1750	GAAGACATCCAGACGCGCTCAGTACCGCTCCATAGAGGTTTTTAATPAGGAGCGGGGTACAGC	1809
DB	1624	GAAGATATTCAACAAGCAATATCGTTCTTGAAGTTCTAATCGGATCTGGCTATAAT	1683
QY	1810	ACCCCTCGGACATCTGGAGACGGCGTGTATGGCATTTGAGCTGGGCAACGGGAGATTAT	1869
DB	1684	ACCCCTGCTGCATTTTGGAGCAGGGATGCATGGCCCTTGAACCTGGGCCACAGGTGACTAT	1743
QY	1870	TTGTTTTCAACCATTTCTGGGGAGACTATTTCCAGAGACGACGACACACATAGCCCATC	1929
DB	1744	TTGTTTGAACCTCATTCAGGGGAAGAGTACACTCGAGATGAAGATCACATTGCATTGATC	1803
QY	1930	ATAGAGCTGCTAGGCAGTATTCCAAGGCACCTTTGCTCTATCTGGAATAATTTCTCGGGAA	1989
DB	1804	ATAGAACTTCTGGGAAGGTGCCTCGCAAGCTCATTTCTGCGAGGAAATATTCCNAGGAA	1863
QY	1990	TTCTTTCAATCGCAGAGAGAACTTGGCAGACATCAACCAAGCTGAAGCCCTGGAGCCTCTTT	2049
DB	1864	TTTTTCACAAAAAAGGTGACCTGAAACATATCACGAAGCTGAAACCTTTGGGGCCTTTTT	1923
QY	2050	GATGTACTTGTGNAAGATTAGCTGGCCCAATGAAGATGCTGCAAGTTTACAGATTTC	2109
DB	1924	GAGGTTCTAGTGGAGAAGTATGATGGTCTCAGGAAGAGGCAGCTGGCTTTCAGAGATTTC	1983
QY	2110	CTGATCCCGATGTTAGAAAAATGGTTCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCTTCGG	2169
DB	1984	TTACTGCCCATGTTGGAGCTGATCCCTGAGAAAGAGAGCCACTGCGCCGAGTGTCTCCGG	2043
QY	2170	CATCCTTGGTTGAATTCCTTA	2189
DB	2044	CACCTTGGCTTAACCTCTA	2063

Search completed: January 23, 2005, 23:48:08
Job time : 1675.02 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2005, 05:33:48 ; Search time 254.768 Seconds
(without alignments)
9075.705 Million cell updates/sec

Title: US-10-799-676-1
Perfect score: 3253
Sequence: 1 tcggcgagcgagtgaggc.....ctataataattactattataa 3253

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfileseq1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3253	100.0	3253	4	US-09-759-359A-1
2	3253	100.0	3253	4	US-10-207-973-1
3	2595.6	79.8	2791	2	US-09-016-000-8
4	666.2	20.5	90541	4	US-09-759-359A-3
5	666.2	20.5	90541	4	US-10-207-973-3
6	632.8	19.5	4299	1	US-08-264-002-1
7	629.4	19.3	923	4	US-09-016-434-703
8	293.4	9.0	1427	3	US-09-173-581-10
9	293.4	9.0	1427	3	US-09-420-915-10
10	284	7.8	793	4	US-09-016-434-79
11	248	7.6	249	4	US-09-016-434-808
12	177.2	5.4	617	3	US-09-385-982-396
13	162.4	5.0	2100	1	US-08-264-002-6
14	159.2	4.9	1731	4	US-09-248-796A-6406
15	117	3.6	569	4	US-09-270-767-1008
16	117	3.6	569	4	US-09-270-767-16290
17	99.6	3.1	7218	1	US-08-232-463-14
18	97.6	3.0	326	4	US-09-513-999C-10592
19	92.8	2.9	641	4	US-09-270-767-1313
20	92.8	2.9	641	4	US-09-270-767-16595
21	62.6	1.9	891	4	US-09-248-796A-11247
22	60.6	1.9	51259	3	US-08-781-891-209
23	60.6	1.9	51259	4	US-09-618-166-209
24	59.8	1.8	1669	3	US-09-461-697-184
25	59.2	1.8	696	3	US-09-461-697-193
26	59.2	1.8	699	3	US-09-461-697-191
27	59.2	1.8	717	3	US-09-461-697-189

28	59.2	1.8	774	3	US-09-461-697-187	Sequence 187, App
29	59.2	1.8	819	3	US-09-461-697-185	Sequence 185, App
30	57.6	1.8	3211	2	US-08-574-959A-8	Sequence 8, Appli
31	57.6	1.8	3211	3	US-09-357-014-8	Sequence 8, Appli
32	57.6	1.8	3901	2	US-08-574-959A-6	Sequence 6, Appli
33	57.6	1.8	3901	3	US-09-357-014-6	Sequence 6, Appli
34	57	1.8	1141	4	US-09-806-708B-22	Sequence 22, Appli
35	55.8	1.7	1434	4	US-09-248-796A-5873	Sequence 5873, Ap
36	55.8	1.7	2409	4	US-09-614-221A-204	Sequence 204, App
37	55.8	1.7	3095	6	5231168-1	Patent No. 5231168
38	55.4	1.7	699	4	US-09-248-796A-9722	Sequence 9722, Ap
39	55.4	1.7	7218	1	US-08-232-463-14	Sequence 14, Appli
40	55	1.7	832	4	US-09-621-976-2813	Sequence 2813, Ap
41	55	1.7	1737	4	US-09-248-796A-581	Sequence 681, App
42	54.8	1.7	478	4	US-09-621-976-10407	Sequence 10407, A
43	54.8	1.7	2223	1	US-08-257-073-4	Sequence 4, Appli
44	54.2	1.7	3337	1	US-08-072-610-1	Sequence 1, Appli
45	54.2	1.7	3337	2	US-08-719-822B-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-759-359A-1
; Sequence 1, Application US/09759359A
; Patent No. 6492153
; GENERAL INFORMATION:
; APPLICANT: ABU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/09/759,359A
; CURRENT FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3253
; TYPE: DNA
; ORGANISM: Human
US-09-759-359A-1

Query Match	100.0%	Score	3253	DB	4	Length	3253
Best Local Similarity	100.0%	Pred. No.	0				
Matches	3253	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
Qy	1	TCGGCGGAGCGAGTGGAGGCTGCAGCCAGCTCGTCTCGGCGCCGCGCGTCCGCGTCCGCGA	60				
Db	1	TCGGCGGAGCGAGTGGAGGCTGCAGCCAGCTCGTCTCGGCGCCGCGCGTCCGCGTCCGCGA	60				
Qy	61	AGCCCCCGCCCGCTTCGCGCGCTCGGAATGAGCTCCGGAAGTGTGGCCATTTCAG	120				
Db	61	AGCCCCCGCCCGCTTCGCGCGCTCGGAATGAGCTCCGGAAGTGTGGCCATTTCAG	120				
Qy	121	GCCGAAAGCGAGGCGGAAAGAGAGAAACATCCGAAAGCGCGAGGCTCAACAGAAA	180				
Db	121	GCCGAAAGCGAGGCGGAAAGAGAGAAACATCCGAAAGCGCGAGGCTCAACAGAAA	180				
Qy	181	GCTCCTTTAGTTCTCTCTCCACCGCCACACACACACACACCGCCACCTTTGCCAGAC	240				
Db	181	GCTCCTTTAGTTCTCTCTCCACCGCCACACACACACACCGCCACCTTTGCCAGAC	240				
Qy	241	CCACACACCGCGAGGAGGAGATCCTGGGATCAGATGATGAGGAGCAAGAGGAC	300				
Db	241	CCACACACCGCGAGGAGGAGATCCTGGGATCAGATGATGAGGAGCAAGAGGAC	300				
Qy	301	CTGCGGACTACTGCAAAAGGTGGATATCATCCAGTGAATAATGGAGACCTCTTTCAATGGC	360				
Db	301	CTGCGGACTACTGCAAAAGGTGGATATCATCCAGTGAATAATGGAGACCTCTTTCAATGGC	360				
Qy	361	CGGTATCATGTTATTAGAAAGCTTGGATGGGGGCACTTCTACTCTCTGGCTGTCTGG	420				
Db	361	CGGTATCATGTTATTAGAAAGCTTGGATGGGGGCACTTCTACTCTCTGGCTGTCTGG	420				

361	CGGTATCATGTTATTAGAAAGCTTGGATGGGGGCATCTTCTCTATGTCGTGCTGTGCTGG	420
421	GATATCGAGGGGAAAAGATTTTGTGCAATGAAAGTTGTAAAAAGTGGCCAGCATTTATACG	480
421		
421	GATATCGAGGGGAAAAGATTTTGTGCAATGAAAGTTGTAAAAAGTGGCCAGCATTTATACG	480
481	GAGACAGCCTTGGATGAAATAAAAATGGCTCAAAATGTTTCGAGAAAGTGATCCCAAGTGAC	540
481	GAGACAGCCTTGGATGAAATAAAAATGGCTCAAAATGTTTCGAGAAAGTGATCCCAAGTGAC	540
541	CCAAAACAAGACATGGTGGTCCAGGCTCATTTGACGACTTCAAGATTTTCAGGCATGAATGGG	600
541	CCAAAACAAGACATGGTGGTCCAGGCTCATTTGACGACTTCAAGATTTTCAGGCATGAATGGG	600
601	ATACATGTCGCATGTCCTTTCGAAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAAA	660
601	ATACATGTCGCATGTCCTTTCGAAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAAA	660
661	TCCAACTATCAAGGCCCTCCAGTACGTTGTTGTAAGAGTATCATTTGACAGTGCCTTCAA	720
661	TCCAACTATCAAGGCCCTCCAGTACGTTGTTGTAAGAGTATCATTTGACAGTGCCTTCAA	720
721	GGTTAGATTACTTTACACAGTAAAGTCAAGATCATTTACTATGACATAAAGCCGGAAT	780
721	GGTTAGATTACTTTACACAGTAAAGTCAAGATCATTTACTATGACATAAAGCCGGAAT	780
781	ATCTTGATGTGTGGATGATGATCATATGTGAGAAAGATGGCAGCTGAGGCCACTGAGTGG	840
781	ATCTTGATGTGTGGATGATGATCATATGTGAGAAAGATGGCAGCTGAGGCCACTGAGTGG	840
841	CAGAAACAGGTGCTCCTCCTCTCTCCAGGTCGTGACGTGAGTACGGCTCCACAGCAGAAA	900
841		
841	CAGAAACAGGTGCTCCTCCTCTCTCCAGGTCGTGACGTGAGTACGGCTCCACAGCAGAAA	900
901	CCTATAGGAAAAATATCTAAAAACAAAAAGAAAAAACTGAAAAAGAAAAACAGAAAGGCGAG	960
901	CCTATAGGAAAAATATCTAAAAACAAAAAGAAAAAACTGAAAAAGAAAAACAGAAAGGCGAG	960
961	GCTCAGTTATTGGAGAGCGCCTCGAGGACATAGAAAGATTTGAGCGCGAAGCTGAAAGG	1020
961	GCTCAGTTATTGGAGAGCGCCTCGAGGACATAGAAAGATTTGAGCGCGAAGCTGAAAGG	1020
1021	AAAAATAATAGAAAAAAACATCACTCAGCTGCACTTCCAAATCAACAGGATGCGCAATAC	1080
1021	AAAAATAATAGAAAAAAACATCACTCAGCTGCACTTCCAAATCAACAGGATGCGCAATAC	1080
1081	TGCCCAGAGGTGAACTTAAAAACAACAGGATTAGAGCGCGCTGAGGCAGAGACTGCA	1140
1081	TGCCCAGAGGTGAACTTAAAAACAACAGGATTAGAGCGCGCTGAGGCAGAGACTGCA	1140
1141	AAGGACAAATGCTGAAGCTTGAGGACCCAGGAGAGAAAAAGATGCTGAGAAAGAAAAATT	1200
1141	AAGGACAAATGCTGAAGCTTGAGGACCCAGGAGAGAAAAAGATGCTGAGAAAGAAAAATT	1200
1201	GAAAAAGATGAAGATGATGTTAGTACAGGAACCTTCGGAACATAGACCCTACGTCGATAGAA	1260
1201	GAAAAAGATGAAGATGATGTTAGTACAGGAACCTTCGGAACATAGACCCTACGTCGATAGAA	1260
1261	TCACCTAAACCAATGGCCATATTGAGAAATGGCCCATTTCTCACTGGAGCAGCAACTGGAC	1320
1261	TCACCTAAACCAATGGCCATATTGAGAAATGGCCCATTTCTCACTGGAGCAGCAACTGGAC	1320
1321	GATGAAGATGATGATGAAGAGACTGCCCAAACTCTGAGGAATATAATCTTTGATGAGCCA	1380
1321	GATGAAGATGATGATGAAGAGACTGCCCAAACTCTGAGGAATATAATCTTTGATGAGCCA	1380
1381	AATGACAGAAAGTATTACATATAGAGCTCCTATGAAACAATTTCAATGGTGAATTCGCCA	1440
1381	AATGACAGAAAGTATTACATATAGAGCTCCTATGAAACAATTTCAATGGTGAATTCGCCA	1440
1441	AATGACAGAAAGTATTACATATAGAGCTCCTATGAAACAATTTCAATGGTGAATTCGCCA	1500
1441	AATGACAGAAAGTATTACATATAGAGCTCCTATGAAACAATTTCAATGGTGAATTCGCCA	1500

QY	1501	GGATCCTTTAGAACCTGTGGCTTCGGCTCTGCTCTCTGAGGGATCACCACTTACTCTGAG	1560
DB	1501	GGATCCTTTAGAAACCTGTGGCCCTGCGGCTCTGTGCTCTCTGAGGGATCACCACTTACTCTGAG	1560
QY	1561	CAAGAGGAGAGCAGTCCATCCATCACAGAAAGCAGAACCGTGTTCAGCCTCCAGTACTGGG	1620
DB	1561	CAAGAGGAGAGCAGTCCATCCATCACAGAAAGCAGAACCGTGTTCAGCCTCCAGTACTGGG	1620
QY	1621	GAATTTGCCAAAAGCAAAAACCCGGGGAGCTGACTTGTGTGGTGAATCCCTCGATCCCGGG	1680
DB	1621	GAATTTGCCAAAAGCAAAAACCCGGGGAGCTGACTTGTGTGGTGAATCCCTCGATCCCGGG	1680
QY	1681	AAATGCAGATAAATTAGAGTGAATAAATTCGTGACCTCGGAAATGCTGTGTGGTGCATAAAA	1740
DB	1681	AAATGCAGATAAATTAGAGTGAATAAATTCGTGACCTCGGAAATGCTGTGTGGTGCATAAAA	1740
QY	1741	CACCTTCACGGAAGACATCCAGACGGGTCAGTACCCGCTCCATAGAGGTTTTAATAGGAGCG	1800
DB	1741	CACCTTCACGGAAGACATCCAGACGGGTCAGTACCCGCTCCATAGAGGTTTTAATAGGAGCG	1800
QY	1801	GGGTACAGCAACCCCTCGGGACATCTGGAGCAGCGGTGTATGGCAATTTGACCTGGCAACG	1860
DB	1801	GGGTACAGCAACCCCTCGGGACATCTGGAGCAGCGGTGTATGGCAATTTGACCTGGCAACG	1860
QY	1861	GGAGATTAATTTGTTTGAACCAACATCTCGGGGAAGACTATTCAGAGACGAAGACCAACATA	1920
DB	1861	GGAGATTAATTTGTTTGAACCAACATCTCGGGGAAGACTATTCAGAGACGAAGACCAACATA	1920
QY	1921	GCCCACTCATAGAGCTGTGAGCAGATATCCNAGGCACCTTGTCTCATCTCGGAATAAT	1980
DB	1921	GCCCACTCATAGAGCTGTGAGCAGATATCCNAGGCACCTTGTCTCATCTCGGAATAAT	1980
QY	1981	TCTCGGGAAATCTTCAATGCCAGAGGAGAACTCGACACATACCAAGCTGAAGAGCCCTGG	2040
DB	1981	TCTCGGGAAATCTTCAATGCCAGAGGAGAACTCGACACATACCAAGCTGAAGAGCCCTGG	2040
QY	2041	AGCCTCTTTTGATGTACTTGTGTGAAAAAGTATGCGTGGCCCCCTGAAGATGCTGCACAGTTT	2100
DB	2041	AGCCTCTTTTGATGTACTTGTGTGAAAAAGTATGCGTGGCCCCCTGAAGATGCTGCACAGTTT	2100
QY	2101	ACAGATTTCTGATCCGATGTTAGAAATGGTTCCAGAAAAACGAGCCCTCAGCTGGCGAA	2160
DB	2101	ACAGATTTCTGATCCGATGTTAGAAATGGTTCCAGAAAAACGAGCCCTCAGCTGGCGAA	2160
QY	2161	TGCGCTTCGGCATCCTTGTTGAAATCTTAGCAAAATCTACCAATATGCAATTCAGACTA	2220
DB	2161	TGCGCTTCGGCATCCTTGTTGAAATCTTAGCAAAATCTACCAATATGCAATTCAGACTA	2220
QY	2221	GCAATATGTTCCAGTACATTTGGACCTAAACCGGTGACTCTCATTTCTTTAAAGGATTAACA	2280
DB	2221	GCAATATGTTCCAGTACATTTGGACCTAAACCGGTGACTCTCATTTCTTTAAAGGATTAACA	2280
QY	2281	GTGAGCTGGCTTCATCTCTCAGACCTTTATTTTGTCTTTGAGGTACTGTGTGTGACATTTT	2340
DB	2281	GTGAGCTGGCTTCATCTCTCAGACCTTTATTTTGTCTTTGAGGTACTGTGTGTGACATTTT	2340
QY	2341	GCTTTTTGTGCACTGTGTGATCTGGGGAAGGGTAGTCTTTTGTCTTCAGCTAAGTAGTTTA	2400
DB	2341	GCTTTTTGTGCACTGTGTGATCTGGGGAAGGGTAGTCTTTTGTCTTCAGCTAAGTAGTTTA	2400
QY	2401	CTGACCAATTTCTTCTGGAAACAATAACNTGCTCTAAGCATTTGTTTCTTGTCTTGTGTG	2460
DB	2401	CTGACCAATTTCTTCTGGAAACAATAACNTGCTCTAAGCATTTGTTTCTTGTCTTGTGTG	2460
QY	2461	ACATTCAAAATGTCATTTTTTTTGAATGAAAAATACTTTTCCCTTTGTGTTTTTGGCAGGTTT	2520
DB	2461	ACATTCAAAATGTCATTTTTTTTGAATGAAAAATACTTTTCCCTTTGTGTTTTTGGCAGGTTT	2520
QY	2521	TGTAACTATTTATGAAGAAATAATTTTACCTCAGTACTATATAATTTACAATCTTAAGAAA	2580
DB	2521	TGTAACTATTTATGAAGAAATAATTTTACCTCAGTACTATATAATTTACAATCTTAAGAAA	2580


```

RESULT 3
US-09-016-000-8
; Sequence 8, Application US/09016000
; Patent No. 5962232
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Lal, Preeti
; APPLICANT: Bandman, Olga
; APPLICANT: Akerblom, Ingrid E.
; APPLICANT: Shah, Purvi
; APPLICANT: Corley, Neil C.
; APPLICANT: Guesler, Karl G.
; TITLE OF INVENTION: PROTEIN KINASE MOLECULES
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016.000
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0465 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2791 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: HEARNOT01
; CLONE: 307624
US-09-016-000-8

Query Match 79.8%; Score 2595.6; DB 2; Length 2791;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2713; Conservative 0; Mismatches 14; Indels 69; Gaps 4

QY 15 GGAGGCTGCAGCCAGCTCGTCTGGCGCCCGCGCTGCGAGGCCCCGAGCCCCCGCCG 74
DB 1 GGAGGCTGCAGCCAGCTCGTCTGGCGCCCGCGCTGCGAGGCCCCGAGCCCCCGCCG 59
QY 75 CTTCCGCGCGCTCGGAATGAGCTCCCGGAAAGTGTCTGCCATTTCAGGCCCGGAAAGCGGAG 134
DB 60 CTTCCGCGCGCTCGGAATGAGCTCCCGGAAAGTGTCTGCCATTTCAGGCCCGGAG 119
QY 135 GCCGAAAGAGAGAAAACATCCGAAAAA-----GCCGAGCCTCAACAGAAAGCTCTTTAG 161
DB 120 GCCGAAAGAGAGAAAACATCCGAAAAAATCAACGAGAGATTGAGCTGCTGATGTGAG 179
QY 162 -----GCCGAGCCTCAACAGAAAGCTCTTTAG 190
DB 180 TTAAGTCTGAGAAGTGTCTCTTTACAGAAAGCGCGAGCCTCAACAGAAAGCTCTTTAG 239

```

QY	1271	CCAATGGCCATATTGAGAAATGCCCATCTTCTACCTGGAGCAGCAAACTGGAGATGAAGATG	1330	Db	2400	CACCTGATCCTGGGAAGGGTAGTCTTTTGTCTTCAGCTAAGTAGTCTTACTGACCATTT	2459
Db	1320	CCAATGGCCATATTGAGAAATGCCCATCTTCTACCTGGAGCAGCAAACTGGAGATGAAGATG	1379	QY	2411	TCTTCTGGAACAATAACATGCTCTTAAGCATTTTCTTGTCTGTGTCGACATTCAAAT	2470
QY	1331	ATGATGAAGAGAGACTGCCCAATCTGAGGAATAATAATCTTGTATGAGCCAAATCGAGAA	1390	Db	2460	TCTTCTGGAACAATAACATGCTCTTAAGCATTTTCTTGTCTGTGTCGACATTCAAAT	2519
Db	1380	ATGATGAAGAGAGACTGCCCAATCTGAGGAATAATAATCTTGTATGAGCCAAATCGAGAA	1439	QY	2471	GTCAATTTTTTGAATGAAAAATACTTTCCCTTTTGTGTTTGGCAGGTTTGTAACTATT	2530
QY	1391	GTGATTAACATATAGCAGCTCTTATGAACAATTCATGGTGAATTTGCCAAATGGACGAC	1450	Db	2520	GTCAATTTTTTGAATGAAAAATACTTTCCCTTTTGTGTTTGGCAGGTTTGTAACTATT	2579
Db	1440	GTGATTAACATATAGCAGCTCTTATGAACAATTCATGGTGAATTTGCCAAATGGACGAC	1499	QY	2531	TATGAAGAAATATTTAGCTGAGTACTATATAATTTTACAATCTTAAGAAATATCAAGTT	2590
QY	1451	ATAAATTTCCCGAGTCAGAGTTCCAGAGTTTCCACCTCGTTGTTCTCTCGATCCTTAG	1510	Db	2580	TATGAAGAAATATTTAGCTGAGTACTATATAATTTTACAATCTTAAGAAATATCAAGTT	2639
Db	1500	ATAAATTTCCCGAGTCAGAGTTCCAGAGTTTCCACCTCGTTGTTCTCTCGATCCTTAG	1559	QY	2591	GGGAACCAAGAAATAGCAAGGAAATGTACAATTTTATCTTCTGGCAAGGGACATCAT	2650
QY	1511	AACTGTGTGGCTCGGGCTCTGTGCTTTCTGAGGGATCAACCTTACTAGCAAGAGGAGA	1570	Db	2640	GGAAACCAAGAAATAGCAAGGAAATGTACAATTTTATCTTCTGGCAAGGGACATCAT	2696
Db	1560	AACTGTGTGGCTCGGGCTCTGTGCTTTCTGAGGGATCAACCTTACTAGCAAGAGGAGA	1619	QY	2651	TCTGTATTATAGTGTATGTAATGCAACCTGTAAATGTTTACTTTGGATTTAAATATGGGA	2710
QY	1571	GCAGTCCATCCCATGACAGAGCAGACGGTTTTCAGCCTCCAGTACTGGGANTTTGCCAA	1630	Db	2697	TCTGTATTATAGTGTATGTAATGCAACCTGTAAATGTTTACTTTGGATTTAAATATGGGA	2756
Db	1620	GCAGTCCATCCCATGACAGAGCAGACGGTTTTCAGCCTCCAGTACTGGGANTTTGCCAA	1679	QY	2711	GGGGGAGCTCAATTTTCAGAAAGCTTAAAAAATAA 2746	
QY	1631	AAGCAAAACCCGGGAGCTGACTTGTGTTGGTGAATCCCTCTGGATCCGCGGAATGCAGATA	1690	Db	2757	-GGGGGAGCTCAATTTTCAGAAAGCTTAAAAAATAA 2791	
Db	1680	AAGCAAAACCCGGGAGCTGACTTGTGTTGGTGAATCCCTCTGGATCCGCGGAATGCAGATA	1739	RESULT 4			
QY	1691	AAATTAGAGTAAAAATTTGCTGACCTGGGAAATGCTTGTGTTGGTGATTAACAACCTCAGGG	1750	US-09-759-359A-3			
Db	1740	AAATTAGAGTAAAAATTTGCTGACCTGGGAAATGCTTGTGTTGGTGATTAACAACCTCAGGG	1799	; Sequence 3, Application US/09759359A			
QY	1751	AGACATCCAGACGGCTGACTACCGCTCCATAGAGTTTATAGAGCGGGGTACAGCA	1810	; Patent No. 6492153			
Db	1800	AGACATCCAGACGGCTGACTACCGCTCCATAGAGTTTATAGAGCGGGGTACAGCA	1859	; GENERAL INFORMATION:			
QY	1811	CCCCTCGGACATCTGGAGCAGGGCTGTATGGCATTTTGAGCTGGCAACGGGAGATTAT	1870	; APPLICANT: ABU-THREIDH, Jane et al			
Db	1860	CCCCTCGGACATCTGGAGCAGGGCTGTATGGCATTTTGAGCTGGCAACGGGAGATTAT	1919	; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC			
QY	1871	TGTTTGAACACATTTCTGGGGAAGACTATTCAGAGACGAACACCATAGCCCATCA	1930	; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES			
Db	1920	TGTTTGAACACATTTCTGGGGAAGACTATTCAGAGACGAACACCATAGCCCATCA	1979	; TITLE OF INVENTION: THEREOF			
QY	1931	TAGAGCTGTAGGACATTTCCAAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGGAAT	1990	; FILE REFERENCE: CLO01043			
Db	1980	TAGAGCTGTAGGACATTTCCAAAGGCACTTTGCTCTATCTGGAATAATTTCTCGGGAAT	2039	; CURRENT APPLICATION NUMBER: US/09/759,359A			
QY	1991	TCTTCAATCGCAGAGGAGAACTGCGACATCAATCAAGCTGAAGCCCTGGAGCTCTTTG	2050	; NUMBER OF SEQ ID NOS: 3			
Db	2040	TCTTCAATCGCAGAGGAGAACTGCGACATCAATCAAGCTGAAGCCCTGGAGCTCTTTG	2099	; SOFTWARE: FastSeq for Windows Version 4.0			
QY	2051	ATGCTATTGGAAGATATGGCTGGCCCATGAAGATGCTGCACAGTTTACAGATTTC	2110	; SEQ ID NO 3			
Db	2100	ATGCTATTGGAAGATATGGCTGGCCCATGAAGATGCTGCACAGTTTACAGATTTC	2159	; LENGTH: 90541			
QY	2111	TGATCCCGATTTAGAAATGGTTCCAGAAAAACGAGCCTCAGCTGGCGAATGCTTCGGC	2170	; TYPE: DNA			
Db	2160	TGATCCCGATTTAGAAATGGTTCCAGAAAAACGAGCCTCAGCTGGCGAATGCTTCGGC	2219	; ORGANISM: Human			
QY	2171	ATCCTTGGTTGAATTTCTTACCAATATTCATTTAGCTAGCTAGCAAAATGTTTC	2230	US-09-759-359A-3			
Db	2220	ATCCTTGGTTGAATTTCTTACCAATATTCATTTAGCTAGCTAGCAAAATGTTTC	2279	Query Match 20.5%; Score 666.2; DB 4; Length 90541;			
QY	2231	CCAGTACATTTGGAACCTTAAACGGTACTCTCATTTTAAACAGGATTTACAGTGGAGTGGC	2290	Best Local Similarity 97.3%; Pred. No. 5.4e-155;			
Db	2280	CCAGTACATTTGGAACCTTAAACGGTACTCTCATTTTAAACAGGATTTACAGTGGAGTGGC	2339	Matches 721; Conservative 0; Mismatches 13; Indels 7; Gaps 4;			
QY	2291	TTCACTCTCAGACCTTTATTTGCTTTGAGGATGCTGTGTTTGAACATTTTGTCTTTTGTG	2350	QY 2004 AGGGAACCTGGACACATCACCAAGCTGAAGCCCTGGAGCCCTTTGATGTTGGA 2063			
Db	2340	TTCACTCTCAGACCTTTATTTGCTTTGAGGATGCTGTGTTTGAACATTTTGTCTTTTGTG	2399	Db 87810 AGGGAACCTGGACACATCACCAAGCTGAAGCCCTGGAGCCCTTTGATGTTGGA 87869			
QY	2351	CACGTGTATCTGGGGAAGGGTAGTCTTTTGTCTTCAGCTAAGTAGTTTACTGACCATTT	2410	QY 2064 AAGATATGGCTGGCCCATGAAGTGTGACAGTTTACAGATTTCCTGATCCCGATGTT 2123			
				Db 87870 AAGATATGGCTGGCCCATGAAGTGTGACAGTTTACAGATTTCCTGATCCCGATGTT 87929			
				QY 2124 AGAAATGTTTCCAGAAAAACGAGCCTCAGCTGGCGAATGCTTGGCAATTCCTTGGTTGAA 2183			
				Db 87930 AGAAATGTTTCCAGAAAAACGAGCCTCAGCTGGCGAATGCTTGGCAATTCCTTGGTTGAA 87989			
				QY 2184 TTCTTAGCAAAATTTACCAATATTCGATTTCTGAGTAGCAAAATGTTCCAGTACATTCGA 2243			
				Db 87990 TTCTTAGCAAAATTTACCAATATTCGATTTCTGAGTAGCAAAATGTTCCAGTACATTCGA 88049			
				QY 2244 CCTAAACGGTACTCTCATTTTAAACAGGATTTACAGTGGAGTGGCTTCATCTCCAGAC 2303			
				Db 88050 CCTAAACGGTACTCTCATTTTAAACAGGATTTACAGTGGAGTGGCTTCATCTCCAGAC 88109			
				QY 2304 CTTTATTTTGTCTTGGGTAAGTGTGTTTGTGATTTGCTTTTGTGCACTGTGATCTCG 2363			

Db 88110 CTTTATTTTGGCTTGGAGTACTGTTGTTTGACATTTTGGCTTTTGTGCACTGTGATCCTG 88169
QY 2364 GGGAGGGTGTCTTTT--TGTCTTCTAGCTAAGTATTTTACTGACCAATTTTCTTT-CTGGAA 2420
Db 88170 GGGAGGGTGTCTTTTGTGCTTCTAGCTAAGTATTTTACTGACCAATTTTCTTCTCTGGAA 88229
QY 2421 ACAATAACATGTCTCTAAGCATTTTCTTGTGTTGTGAGACATTTCAATGTCAATTTT 2480
Db 88230 ACAATAACATGTCTCTAAGCATTTTCTTGTGTTGTGAGACATTTCAATGTCAATTTT 88289
QY 2481 TGAATGAAAAATACATTTTCCCTTTTGGCAGGTTTGTAACTATTATTAAGAA 2540
Db 88290 TGAATGAAAAATACATTTTCCCTTTTGGCAGGTTTGTAACTATTATTAAGAA 88349
QY 2541 TATTTTGTAGTGTACTATTAATTTTCAATCTTAAAGAAATTTATCAAGTTGGGAACCAAG 2600
Db 88350 TATTTTGTAGTGTACTATTAATTTTCAATCTTAAAGAAATTTATCAAGTTGGGAACCAAG 88409
QY 2601 AAAATAGCAAGGAAATGACAAATTTTCTTCTGCAAGGACATCATTTCTGTATTA 2660
Db 88410 AATA---GCAAGAAATGACAAATTTTCTTCTGCAAGGACATCATTTCTGTATTA 88466
QY 2661 TAGTGTATGTAATGCACCTGTAAATTTTACTTTTGGATTAATATGGAGGGGACTC 2720
Db 88467 TAGTGTATGTAATGCACCTGTAAATTTTACTTTTCCATTAATATGGA-GGGGACTC 88525
QY 2721 AAATTTTCAGAAAAGCTAAAAA 2741
Db 88526 AAATTTTCAGAAAAGCTAACCA 88546

RESULT 5
US-10-207-973-3
; Sequence 3, Application US/10207973
; Patent No. 6753175
; GENERAL INFORMATION:
; APPLICANT: ASU-THREIDEH, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001043
; CURRENT APPLICATION NUMBER: US/10/207,973
; CURRENT FILING DATE: 2002-07-31
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; TYPE: DNA
; ORGANISM: Human
US-10-207-973-3

Query Match 20.5%; Score 666.2; DB 4; Length 90541;
Best Local Similarity 97.3%; Pred. No. 5.4e-155;
Matches 721; Conservative 0; Mismatches 13; Indels 7; Gaps 4;
QY 2004 AGGAGAACTGCGACACATCACCAGCTGAAGCCCTTGTGATGTACTTGTGGA 2063
Db 87810 AGGAGAACTGCGACACATCACCAGCTGAAGCCCTTGTGATGTACTTGTGGA 87869
QY 2064 AAAGTATGCTGCCCCCATGAAGTCTGCACAGTTTACAGATTTTCCGATCCGATGTT 2123
Db 87870 AAAGTATGCTGCCCCCATGAAGTCTGCACAGTTTACAGATTTTCCGATCCGATGTT 87929
QY 2124 AGAAATGGTTCCAGAAAAACGAGCCCTCAGCTGCGAATCCCTTGGCATCTTGGTTGAA 2183
Db 87930 AGAAATGGTTCCAGAAAAACGAGCCCTCAGCTGCGAATCCCTTGGCATCTTGGTTGAA 87989
QY 2184 TTCTTAGCAAAATCTACCAATATTTGCAATCTGAGCTAGCAAAATGTTCCAGTACATTTGGA 2243
Db 87990 TTCTTAGCAAAATCTACCAATATTTGCAATCTGAGCTAGCAAAATGTTCCAGTACATTTGGA 88049
QY 2244 CCTAAACGGTGTACTCTCAATTTCTTTAAACAGGATTAACAGTGTGCTGCTTCATCCTCAGAC 2303

Db 88050 CTTAAACGGTGTACTCTCAATTTTAAACAGGATTAACAGTGTGCTGCTTCATCCTCAGAC 88109
QY 2304 CTTTATTTTGTCTTGGAGTACTGTTGTTTGAATTTTGTCTTTTGTGCACTGTGATCCTG 2363
Db 88110 CTTTATTTTGTCTTGGAGTACTGTTGTTTGAATTTTGTCTTTTGTGCACTGTGATCCTG 88169
QY 2364 GGGAGGGTGTCTTTT--TGTCTTCTAGCTAAGTATTTTACTGACCAATTTTCTTT-CTGGAA 2420
Db 88170 GGGAGGGTGTCTTTTGTGCTTCTAGCTAAGTATTTTACTGACCAATTTTCTTCTCTGGAA 88229
QY 2421 ACAATAACATGTCTCTAAGCATTTTCTTGTGTTGTGAGACATTTCAATGTCAATTTT 2480
Db 88230 ACAATAACATGTCTCTAAGCATTTTCTTGTGTTGTGAGACATTTCAATGTCAATTTT 88289
QY 2481 TGAATGAAAAATACATTTTCCCTTTTGGCAGGTTTGTAACTATTATTAAGAA 2540
Db 88290 TGAATGAAAAATACATTTTCCCTTTTGGCAGGTTTGTAACTATTATTAAGAA 88349
QY 2541 TATTTTGTAGTGTACTATTAATTTTCAATCTTAAAGAAATTTATCAAGTTGGGAACCAAG 2600
Db 88350 TATTTTGTAGTGTACTATTAATTTTCAATCTTAAAGAAATTTATCAAGTTGGGAACCAAG 88409
QY 2601 AAAATAGCAAGGAAATGACAAATTTTCTTCTGCAAGGACATCATTTCTGTATTA 2660
Db 88410 AATA---GCAAGAAATGACAAATTTTCTTCTGCAAGGACATCATTTCTGTATTA 88466
QY 2661 TAGTGTATGTAATGCACCTGTAAATTTTACTTTTGGATTAATATGGAGGGGACTC 2720
Db 88467 TAGTGTATGTAATGCACCTGTAAATTTTACTTTTCCATTAATATGGA-GGGGACTC 88525
QY 2721 AAATTTTCAGAAAAGCTAAAAA 2741
Db 88526 AAATTTTCAGAAAAGCTAACCA 88546

RESULT 6
US-08-264-002-1
; Sequence 1, Application US/08264002
; Patent No. 5859019
; GENERAL INFORMATION:
; APPLICANT: GUI, JIAN-PANG
; APPLICANT: FU, XIANG-DONG
; TITLE OF INVENTION: NOVEL PROTEIN SERINE KINASE, SRPK1
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 Century Park East, Fifth Floor
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/264,002
; FILING DATE: 22-JUN-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: TUMARKIN PH.D., LISA A.
; REGISTRATION NUMBER: P-38,347
; REFERENCE/DOCKET NUMBER: PD3590
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4299 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
IMMEDIATE SOURCE:
CLONE: SRPK1
FEATURE:
NAME/KEY: CDS
LOCATION: 109..2073
US-08-264-002-1

Query Match 19.5%; Score 632.8; DB 1; Length 4299;
Best Local Similarity 60.4%; Pred. No. 2e-147;
Matches 1172; Conservative 0; Mismatches 672; Indels 96; Gaps 4;

QY 250 CCGAGCCAGAGAGAGATCTCGGATCAGATGATGAGGAGGAGGAGCCCTGCGGAC 309
DB 232 CCAGAGCAGGAGAGAGATCTCGGATCTGATGATGATGAGGAGGAGATCTTAATGAT 291

QY 310 TACTGCAAAAGGTGGATATCATCCAGTGAATAATGGAGACCTCTTCAATGGCCGGTATCAT 369
DB 292 TATGTGAAGAGGTTATCATCTTGTGAAAATGGAGATCTATTCATGGGAGATACCAT 351

QY 370 GTTATTAGAAAGCTTGGATGGGGGCACTTCTCTA CTGTCTGCTGTGCTGGGATATGCG 429
DB 352 GTGATCCGAAAGTTAGGCTGGGACACATTTTCAACAGTATGGTTATCATGGGATATTCAG 411

QY 430 GGGAAAGATTTGTCATGAATGAAGTTGAAAAGTGGCCAGCATTTATACGGAGAGGCC 489
DB 412 GGGAAAGATTTGTCGCAATGAAGTAGTAAAGTGTGAACATTAACATGAAACAGCA 471

QY 490 TTGGATGAATAAAATTTGCTCAAAATGTTTCGAGAAAGTATCCAGTGAACCCAAACAAA 549
DB 472 CTAGATGAAATCCGTTGCTGAAGTCAGTTTCGCAATTCAGACCTTAATGATCCAAATAGA 531

QY 550 GACATGGTGTCCAGCTCATTTAGCGACTTCAGATTTTCAGGCGATGATGGGATACATGTC 609
DB 532 GAAATGGTGTTCACACTACTAGATGACTTTAAATAATACGAGGTTAATGGAACACATATC 591

QY 610 TGCATGTCTTTCGAAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAAAATCCAACTAT 669
DB 592 TGCATGTATTTGAAGTTTGGGGCATCTCTGCTCAAGTGGATCATCAAAATCCAACTAT 651

QY 670 CAAGGCCCTCCAGTACGTTGTGTGAAGATATCAATTCGACAGGTCCTTCAAGGGTTAGAT 729
DB 652 CAGGGCTTCCACTGCTGCTGTCAAAAATAATTAATTCAGCAGTGTGTACAGGCTTTGAT 711

QY 730 TACTTACACAGTAAGTGCNAGATCATTTCACTGACATAAAGCCGGAATAATCTTTGATG 789
DB 712 TATTTACATACCAAGTCCCGTATCATCCACATGACATTAACACAGAACATCTTATTG 771

QY 790 TGTGTGATGATGCATATGTGAAGAAATGGCAGCTGAGGCCACTGAGTGGCGAAGCA 849
DB 772 TCAGTGAATGAGCAGTACATTCGGAGGCTGGCTGCAGAAAGCAACAGAAATGGCAGCATCT 831

QY 850 GGTGCTCTCTCTCTCAGGCTGTCAGTGTGAGTACGGCTCCACAGCAGAAACCTATAGGA 909
DB 832 GGAGCTCTCTCGGCTTCCGATCTGCACTGAGTCTGCTCTCCAGCCTTAACAGCTGAC 891

QY 910 AAAATATCTTAAACAAAAGAAAACCTGAAAAGAAAACAGAGAGGCGAGGCTGAGTTA 969
DB 892 AAAATGTCAAAGATAGAGAGAAATTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 951

QY 970 TTGGAGAGCCGCTCGAGGAGATAGAGAAATTTGAGGCGAGAGCTGAGAGAGAGAGAGAG 1029
DB 952 CTAG 997

QY 1030 GAAGAAACATCACTCAGCTGCATCTTCCAAATGACAGGATGGCGAATATCTGCGCAGAG 1089
DB 998 -----GCCCTGGCAG 1020

QY 1090 GTGAAACTAAAACAAACAGGATTTAGAGAGCGGCTGAGGCGAGAGACTGCAAGAGCAAT 1149
DB 1021 AAGCAAG 1080

QY 1150 GGTGAAGCTGAGGACCAAGAGAGAGAGAGAGATGCTGAGAGAGAGAGAGAGAGAGAGAT 1209
DB 1081 ATGACCCAAAGAAAACCTTGAGAGTCAAGTACCAATTTGGCCAGGATCAACACGCTTATGGAA 1140

QY 1210 GAAGATGATGTAGATCAGGAACTTTCGGAACATAGACCCCTACGCTGATAGAGATCACTAAA 1269
DB 1141 CGTGATACAGAGGGTGTGAGCAGAGAAATTAATTGCAATGGAGTGAATGAAGTCATTAT 1200

QY 1270 ACCAATGGCCATATTTCAGAAATGGCCCATTTCTCACTGGAGCAGCAACTGGAACGATGAAGAT 1329
DB 1201 TATACTCAGNACAGTAAATTAATGAACATTTGAGACATAGAGAGAGATCTCATATATGCTAAT 1260

QY 1330 GATGATGAAGAGACTGCCCCAAATCTCTGAGGAAATATAATCTTTGATGAGCCAAATGCAGAA 1389
DB 1261 GACTGTGTGTCCTCAAAATTTGAATCAGGAATCTAGTTTCTTAAGTCTTCCAAATGGAGA- 1319

QY 1390 AGTGATTAACATATAGCAGCTCTATGAACAAATTCATATGGTGAATTTGCAAAATGGACGA 1449
DB 1320 -----CAGCAGCAGATCTCAAGAAACAGACTCTTTGTACACCTATAACA 1362

QY 1450 CATAAAATTTCCGAGTCAAGTTCCAGAGTTTTTCCACTCTGTTGTTCTCTGGATCCTTA 1509
DB 1363 TCTGAGGTGTGACACACCATGGTGTGCCAGTCTTCTCTCAACTGATAGGTCACTATTCAGT 1422

QY 1510 GAACCTGTGGCTCGGCTCTGTGCTTTCTGAGGGATCACTTACTGAGCAGCAAGAGGAG 1569
DB 1423 GAAC-----AACACATTTAGCCAACTTCAAGAAAGCATTCGGGC 1460

QY 1570 AGCAGTCCATCCCATGACAGAGCAAGACGGTTTCAGCCCTCCAGTACTGGGGATTTGCCA 1629
DB 1461 AGAGTACCTCTGTGAGATGAACAGAGCAAGAAACATTAACGGACCACTGGACA-----AC 1515

QY 1630 AAAGCAAAACCCGGGAGCTGACTTGTGTGTAATCCCTGATCCGCGGAATGCAGAT 1689
DB 1516 AAAGGAAATCCAGGCTGGAAATTTTCTTGTAAATCCCTTGGCCGCAACAACTTCACT 1635

QY 1690 AAAATTAGTATAAAATTTGCTGACCTGGGAAATGCTTTGTTGGGTGCAATAAACACTTCACG 1749
DB 1576 AAGCTCAAGTGAAGATTTGCTGACCTTGGAAATGCTTTGTTGGGTGCAACAACTTCACT 1635

QY 1750 GAAGACATCCAGACCGCTCAGTACCGCTCCATAGAGGTTTTTAATAGAGCGGGGTACAGC 1809
DB 1636 GAAGATATCAAAACAGGCAATATCTTCTTGGAAATTTCTAATCGGATCTGGCTATAAT 1695

QY 1810 ACCCTCGGACATCTGGACACGCGCTGTATGCAATTTGAGCTGGCAACGGGAGATAT 1869
DB 1696 ACCCTGCTGACATTTGGAGCACCGCATGCAATGGCTTTGNACTGGCCACAGGTGACTAT 1755

QY 1870 TTGTTTGAACACATTTCTGGGGAAGACTATTTCCAGACACGAAAGCAACATAGCCCATC 1929
DB 1756 TTGTTTGAACCTCATTCAGGGGAGAGGTACACTCGAGATGAAGATCACTTGCATTTGATC 1815

QY 1930 ATAGAGCTGTAGGAGTATTTCCAGGCACTTTGCTCTATCTGGAATAATTTCTCGGAA 1989
DB 1816 ATAGAACTTCTGGGGAAGGCTCTCGCAAGCTCATTTGTGGCAGGAAATAATTTCCAGGAA 1875

QY 1990 TTCTTCAATCGCAGAGGAGAACTGCGACACATCACTCAAGCTGAAAGCCCTGGAGCTCTTT 2049
DB 1876 TTTTTCACCAAAAAGGTGACCTGNAACATATCAGAGCTGAAACCTTTGGGGCTTTTT 1935

QY 2050 GATGTATTTGTGAAAAAGTATGGCTGGCCCATGAGATGCTGCACAGTTTACAGATTC 2109
DB 1936 GAGGTTCTAGTGGAGAAATGATGAGTGTCTCAGGAAAGGAGGCTGGCTTACAGATTC 1995

QY 2110 CTGATCCCGATGTTAGAAATGTTCCAGAAAAACGAGCTCAGCTGGCGGAATTCCTTCGG 2169
DB 1996 TTACTGCCCATGTTGAGCTGATCTCTGAGAGAGAGCCACTGCGCGGAGTGTCTCCGG 2055

QY 2170 CATCTTGGTTGAATTTCTTA 2189
DB 2056 CACCCTTGGCTTAACCTCTTA 2075


```

1  RESULT 7
2  US-09-016-434-703
3  ; Sequence 703, Application US/09016434
4  ; Patent No. 6500938
5  ; GENERAL INFORMATION:
6  ; APPLICANT: Janice Au-Young
7  ; APPLICANT: Jeffrey J. Seilhamer
8  ; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
9  ; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
10 ; NUMBER OF SEQUENCES: 1490
11 ; CORRESPONDENCE ADDRESS:
12 ; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
13 ; STREET: 3174 PORTER DRIVE
14 ; CITY: PALO ALTO
15 ; STATE: CALIFORNIA
16 ; COUNTRY: USA
17 ; ZIP: 94304
18 ; COMPUTER READABLE FORM:
19 ; MEDIUM TYPE: Floppy disk
20 ; COMPUTER: IBM PC compatible
21 ; OPERATING SYSTEM: PC-DOS/MS-DOS
22 ; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
23 ; CURRENT APPLICATION DATA:
24 ; APPLICATION NUMBER: US/09/016,434
25 ; FILING DATE: HERewith
26 ; CLASSIFICATION:
27 ; PRIOR APPLICATION DATA:
28 ; APPLICATION NUMBER:
29 ; FILING DATE:
30 ; CLASSIFICATION:
31 ; ATTORNEY/AGENT INFORMATION:
32 ; NAME: Zeller, Karen J.
33 ; REGISTRATION NUMBER: 37,071
34 ; REFERENCE/DOCKET NUMBER: PA-0002 US
35 ; TELECOMMUNICATION INFORMATION:
36 ; TELEPHONE: (650) 855-0555
37 ; TELEFAX: (650) 845-4166
38 ; INFORMATION FOR SEQ ID NO: 703:
39 ; SEQUENCE CHARACTERISTICS:
40 ; LENGTH: 923 base,pairs
41 ; TYPE: nucleic acid
42 ; STRANDEDNESS: single
43 ; TOPOLOGY: linear
44 ; IMMEDIATE SOURCE:
45 ; LIBRARY: HEARNOT01
46 ; CLONE: 307624
47 ;
48 ; US-09-016-434-703

```

Query Match	19.3%;	Score 629.4;	DB 4;	Length 923;
Best Local Similarity	88.3%;	Pred. No. 5.7e-147;		
Matches 700;	Conservative 0;	Mismatches 89;	Indels 4;	Gaps 3;
Qy	157	AAAAGCGGAGCTCAACAGAAAGCTCTT	TAGTTCTCTCTCTCCACCGCCACCA	216
Db	43	AAAAGGGGCGCCCTCGCCAAAAGCTCAN	TTTTTCTCTNTTCTCTCCACCGCCACCA	102
Qy	217	CCACCACCGCACCTTTTGCCAGACCCCA	CACCCCGGAGCCAGAGGAGGAGATCCT	276
Db	103	CCACCACCGCNCCTTTTCCNGACCCCA	CACCCCGGAGCCAGAGGAGGAGATCCT	162
Qy	277	TCAGATGATGAGGAGCAAGAGGACCCCT	GCGGACTACTGCAAGGTGGATATCATCC	336
Db	163	TCAGATGATGAGGAGCAAGAGGACCCCT	GCGGACTACTNCNANTGTGGATATCATCC	222
Qy	337	AAAATTGGAGACCTCTTCAATGGCGCGT	ATCATGTTATTAGAAAGCTTGGATGGGGC	396
Db	223	AAAATTGGAGACCTCTTCAATGGCGCGT	ATCATGTTATTAGAAAGCTTGGATGGGGC	282
Qy	397	TTCTCTACTGTCTCGCTGTGCTGGGATAT	GCAGGGGAAAAGATTTGTTGCAATGAAAG	456
Db	283	TTCTCTACTGTCTCGGCTGTGCTGGGATAT	GCAGGGGAAAAGATTTGTTGCAATGAAAG	342
Qy	457	GTA AAAAGTGC CCAGCACTTATACGGGAG	CAGCCCTTGGATGAAATAAAATTCGCTCA	516

Db 860 GCAGATAAGATCAAGATCGAGACCTGGGCAACGCTGCTGGTGCAACAGCAC 919
Qy 1744 TTCACGGAAGACATCCAGACGCTGACGCTCCATAGAGGTTTAAATAGGAGCGGG 1803
Db 920 TTCACGGAAGACATCCAGACGCTGACGCTCCATAGAGGTTTAAATAGGAGCGGG 979
Qy 1804 TACAGACCCCTCGGGAATCTGGAGCAGCGGCTGATGGCAATTTGAGCTGGCAACGGGA 1863
Db 980 TACGCCCCCGGCGAGACATCTGGAGCAGCGCTGATGGCTTTCGAGCTGGCCACTGGT 1039
Qy 1864 GATTATTTTGTGACCACTTCTGGGAGCACTTCCAGAGCAAGAGACCAATAGCC 1923
Db 1040 GACTACCTGTTTCGAGCGCATCTCGAGAGCACTTCCAGAGCAAGAGACCAATAGCC 1099
Qy 1924 CACATCATAGAGCTGCTAGGAGTATTCAGAGGCACTTTGCTCTATCTCGAAAAATATTCT 1983
Db 1100 CACATAGTGGAGCTTCTGGGGAGCATCCCGGAGCTTCCGCTCTCAGGCGGCTATTTC 1159
Qy 1984 CGGGAATTTCTCAATCGCAGAGGAGAACTGCGACACATCACCAGCTGAAGCCCTGGAGC 2043
Db 1160 CGGGAGTCTTCAACCGGAGAGGAGCTGCGGCACTCCACATCTCAAGCACTGGGGC 1219
Qy 2044 CTCTTTGATGACTTGTGAAAAAGTATGGCTGGCCCCCATGAGATGCTGCAAGTTTACA 2103
Db 1220 CTGTACGAGTACTCATGAAAGTACGAGTGGCCCCCTAGAGCAGGCCACACAGTTTACG 1279
Qy 2104 GATTTCCTGATCCGATGTTAGAAATGTTCCAGAAAAACGAGCTCAGCTGGCGAATGC 2163
Db 1280 GCCTTTCTGCTGCCATGAATGATGATACATCCCGGAAAAAGCGGCGCAGTGCGGCTGACTGC 1339
Qy 2164 CTTCGGCATCCTGGTTGAAATCTTAG 2190
Db 1340 CTCAGCAGCCCTGGCTCCAAACCTAG 1366

RESULT 9

US-09-420-915-10
; Sequence 10, Application US/09420915
; Patent No. 6264947
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yalda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PR-0614 US
; CURRENT APPLICATION NUMBER: US/09/420,915
; EARLIER FILING DATE: 1999-10-20
; EARLIER FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 10
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 119819
US-09-420-915-10

Query Match 9.0%; Score 293.4; DB 3; Length 1427;
Best Local Similarity 69.8%; Pred. No. 4.3e-63;
Matches 396; Conservative 0; Mismatches 171; Indels 0; Gaps 0;
Qy 1624 TTGCCAAAGAAAAACCGGCGAGCTGACTTGTGTTGATCCCTGGATCCCGGAAT 1683
Db 800 TCGCCTAGCACACCATTCGGTGCTCGAACCTCTCTGGTGAACCCCTGGAGCCCAAAAT 859

Qy 1684 GCAGATAAAATAGAGTAAAAATTTGCTGACCTGGGAAATGCTTGTGGTGCAATAAAC 1743
Db 860 GCAGATAAGATCAAGATCAAGATCGAGACCTGGGCAACGCTGCTGGTGCAACAGCAC 919
Qy 1744 TTCACGGAAGACATCCAGACGCTGACGCTCCATAGAGGTTTAAATAGGAGCGGG 1803
Db 920 TTCACGGAAGACATCCAGACCTCGGAGTACCGGGCGCTGAGGTGCTGATCGGCCCGGA 979
Qy 1804 TACAGACCCCTCGGAGACATCTGGAGCAGCGGCTGATGGCAATTTGAGCTGGCAACGGGA 1863
Db 980 TACGCCCCCGGCGAGACATCTGGAGCAAGAGCTGATGGCTTTCGAGCTGGCCACTGGT 1039
Qy 1864 GATTATTTTGTGAAACCACTTCTGGGGAAGACTTATTCAGAGCAAGAGACCAATAGCC 1923
Db 1040 GACTACCTGTTTCGAGCGCATCTCGAGAGCACTTACAGTCTGATGAGGACCAATCGCT 1099
Qy 1924 CACATCATAGAGCTGCTAGGAGTATTCAGAGGCACTTTGCTCTATCTCGAAAAATATTCT 1983
Db 1100 CACATAGTGGAGCTTCTGGGGAGCATCCCGGAGCTTCCGCTCTCAGGCGGCTATTTC 1159
Qy 1984 CGGGAATTTCTCAATCGCAGAGGAGAACTGCGACACATCACCAGCTGAAGCCCTGGAGC 2043
Db 1160 CGGGAGTCTTCAACCGGAGAGGAGCTGCGGCACTCCACATCTCAAGCACTGGGGC 1219
Qy 2044 CTCTTTGATGACTTGTGAAAAAGTATGGCTGGCCCCCATGAGATGCTGCAAGTTTACA 2103
Db 1220 CTGTACGAGTACTCATGAAAGTACGAGTGGCCCCCTAGAGCAGGCCACACAGTTTACG 1279
Qy 2104 GATTTCCTGATCCGATGTTAGAAATGTTCCAGAAAAACGAGCTCAGCTGGCGAATGC 2163
Db 1280 GCCTTTCTGCTGCCATGAATGATGATACATCCCGGAAAAAGCGGCGCAGTGCGGCTGACTGC 1339
Qy 2164 CTTCGGCATCCTGGTTGAAATCTTAG 2190
Db 1340 CTCAGCAGCCCTGGCTCCAAACCTAG 1366

RESULT 10

US-09-016-434-79
; Sequence 79, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:

; SEQ ID NO 6406
; LENGTH: 1731
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-6406

Query Match 4.9%; Score 159.2; DB 4; Length 1731;
Best Local Similarity 55.7%; Pred. No. 1.4e-29;
Matches 305; Conservative 0; Mismatches 243; Indels 0; Gaps 0;
QY 1682 ATGCAGATAAAATTAGAGTAAATAATTCGTGACCTGGGAAATGCTTGTGGGTGCATPAAC 1741
Db |||||
QY 980 ATAATGAATTAATATCTGTAAATTCGTGTAATGCTTGGTAATGCTTGGACAAATCATC 1039
Db |||||
QY 1742 ACTTCACGGAGACATCCAGACCGCTCAGTACCGCTCCATAGAGGTTTAAATAGGAGCGG 1801
Db |||||
QY 1040 ATTTTACTGATGAAATCCAAACAGACAATATCGAGCCCCCGAAATTTTAAATTTGGTTATT 1099
Db |||||
QY 1802 GGTACAGACCCCTCGCGGACATCTGGAGCACGCGGTGTATGGCAATTTGAGCTGGCAACGG 1861
Db |||||
QY 1100 ATTTGGGTGCCCTGAGTGATTTATGTCATTTGCTGTTTAAATTTGAATTTAACTG 1159
Db |||||
QY 1862 GAGATTATTTGTTGAACCAATCTCTGGGGAAGACTATTCCAGAGACGAGACCAACATAG 1921
Db |||||
QY 1160 GTGATTATCTATTTGATCCCTAGAGATGGTAAATCATATATAAAAGATGATCATATTG 1219
Db |||||
QY 1922 CCCACATCATAGAGCTGCTAGGACGATTTCCAGGCACTTTGCTCTATCTGGHAAATATT 1981
Db |||||
QY 1220 CCCAAATTTGTAATTTGATTTGGACCAATTTCTTAATCAAAATGTTAAAGAAAGTTATTATG 1279
Db |||||
QY 1982 CTCGGGAATTTCTCAATCGCAGAGGAGAACTGGCACACATCACCAAGCTGAAGCCCTGGA 2041
Db |||||
QY 1280 CTCGAGAAATTTTCAATTTCAAGTATGAATTAAGAAGATCATGAAATTAACCTTTGGG 1339
Db |||||
QY 2042 GCCTCTTTGATGACTTTGTGGAAGTATGGCTGGCCCAATGAAGATGCTGCACAGTTTA 2101
Db |||||
QY 1340 GGTTACAAGATGATTAAATTTGAAATAATATAAATTCCTTTGAATGATGCTATTGAAATTA 1399
Db |||||
QY 2102 CAGATTTCTGTATCCCGATGTTAGAAATGGTTCCAGAAAAACGAGCCTCAGCTGGCGAAT 2161
Db |||||
QY 1400 GTGAATTTTTTATTTGCCAATTTGTAATTTAAACCCAGAGAAAGAGCTGCTGGAGGAA 1459
Db |||||
QY 2162 GCCTTCGGCATCTCTGGTTGAAATTTCTAGCAATTTCTACCAATTTGCAATTTCTGAGCTAG 2221
Db |||||
QY 1460 TGTGTAATCATCTCTGTTTGAAGATGCTTTAGTTTGAATAATGTGGTTTGGAAACGAC 1519
Db |||||
QY 2222 CAAATGTT 2229
Db |||||
QY 1520 CAGTTGGT 1527
Db |||||

RESULT 15
US-09-270-767-1008
; Sequence 1008, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1008
; LENGTH: 569
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-1008

Query Match 3.6%; Score 117; DB 4; Length 569;
Best Local Similarity 67.3%; Pred. No. 2.4e-19;
Matches 165; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

QY 1690 AAAATTAGAGTAAAAATTCGTGACCTGGGAAATGCTTGTGGGTGCATAAACACTTCAAG 1749
Db |||||
QY 324 AACGTTTCGGGTAAAGATCGCCGACTTGGGAAACGCTGTACGACTACCATCACTTTACT 383
Db |||||
QY 1750 GAAGACATCCAGACGCGTCCAGTACCGCTCCATAGAGGTTTAAATAGGAGCGGGGTACAGC 1809
Db |||||
QY 384 GAGGACATTCAGACTCGCCAGTATCGATCAATCGAGGTTCTTTTGGGAGCGCGTACAAT 443
Db |||||
QY 1810 ACCCTTCGGGACATCTGGAGCAGCGGTGTATGGCAATTTGAGCTGGCAACGGGAGATTAT 1869
Db |||||
QY 444 TATACCGCGGACATCTGGAGCACAGCTGTTTGGCCTTCGAGCTGGCCACCGCGACTAC 503
Db |||||
QY 1870 TTGTTTGAACCAATTTCTCGGGAAGACTATTCCAGAGACGAGACCAACATAGCCACATC 1929
Db |||||
QY 504 CTGTTGACCCCTCACGCCGAGAGCTTTACGTCGGGACGAGACCCTTGGCGCATATT 563
Db |||||
QY 1930 ATAGA 1934
Db |||||
QY 564 GTGGA 568
Db |||||

Search completed: January 23, 2005, 11:11:32
Job time : 258.768 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2005, 03:34:58 ; Search time 1468.41 Seconds
(without alignments)
11629.152 Million cell updates/sec

Title: US-10-799-676-1
Perfect score: 3253
Sequence: 1 tcggcgagcagtgaggc.....ctatataattactattataa 3253

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues
Total number of hits satisfying chosen parameters: 8269772
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N Geneseq 23Sep04:*
- 1: geneseqn1980s:*
 - 2: geneseqn1990s:*
 - 3: geneseqn2000s:*
 - 4: geneseqn2001as:*
 - 5: geneseqn2001bs:*
 - 6: geneseqn2002as:*
 - 7: geneseqn2002bs:*
 - 8: geneseqn2003as:*
 - 9: geneseqn2003bs:*
 - 10: geneseqn2003cs:*
 - 11: geneseqn2003ds:*
 - 12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3253	100.0	3253	10	ADJ37688 Human kin
2	3249.8	99.9	3253	6	AB52846 cDNA enco
3	2595.6	79.8	2791	2	AAX89851 Human pro
4	2533.8	77.9	3702	8	AAL54214 SR protei
5	2533.8	77.9	3715	12	ADJ96579 Human SRP
6	2497	76.8	3745	6	ABK51173 cDNA enco
7	2497	76.8	3745	8	AAD53129 Human DNA
8	2497	76.8	3745	8	AAL54215 SR protei
9	2497	76.8	3745	12	ADO19346 Human PRO
10	2039.6	62.7	2094	8	AAL54217 SR protei
11	1753.6	53.9	1956	8	AAL54216 SR protei
12	1237	38.0	3213	5	AAS66342 DNA enco
13	712.8	21.9	1005	12	ADN98453 Novel hum
14	712.8	21.9	1005	12	ADO00022 Novel hum
15	709.6	21.8	1017	8	AAL54219 SR protei
16	689.2	21.2	113033	8	AAL54213 SR protei
17	666.2	20.5	90541	10	ADJ37690 Human kin
18	664.6	20.4	90541	6	AB52847 Human SR
19	636.4	19.6	4392	5	AAS78896 DNA enco
20	632.8	19.5	4326	6	ABK84213 Human cDN
21	632.8	19.5	4326	6	ABN96797 Gene #329

22	632.8	19.5	4326	8	AAD53128 Human DNA
23	632.8	19.5	4326	8	AAL54209 SR protei
24	632.8	19.5	4346	2	AAT45897 Serine ki
25	632.8	19.5	4760	8	AAL54211 SR protei
26	629.6	19.4	1992	8	AAL54212 SR protei
27	629.6	19.4	4349	10	ADE54027 Human ova
28	629.6	19.4	4639	5	ADL63775 Human ova
29	629.4	19.3	923	10	ACA56105 Human big
30	629.4	19.3	923	12	ADI55901 Human pol
31	613.6	18.9	4244	8	AAL54210 SR protei
32	600	18.4	600	8	ABZ19725 Group III
33	488.6	15.0	574	12	ACH76482 Human gen
34	466	14.3	466	4	ABA70433 Human fo
35	466	14.3	466	4	AAT50585 Probe #19
36	466	14.3	466	4	ABA37079 Probe #15
37	466	14.3	466	4	AAX44603 Human bon
38	466	14.3	466	4	AAL18673 Human bra
39	466	14.3	466	4	ABS44262 Human liv
40	466	14.3	466	6	ABS18841 Human gen
41	441.8	13.6	533	12	ACH79758 Human gen
42	441.8	13.6	533	12	ACH68330 Human gen
43	404.2	12.4	1602	4	AAF44709 Novel pro
44	404.2	12.4	1602	12	ADI29407 Human MAR
45	404.2	12.4	1835	8	AAL54218 SR protei

ALIGNMENTS

RESULT 1
ADJ37688
ID ADJ37688 standard; cDNA; 3253 BP.
XX
AC ADJ37688;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human kinase cDNA.
XX
KW Kinase protein; drug screening assay; gene therapy;
KW pharmacogenomic analysis; testis; nervous tissue; fetal; lung;
KW brain anaplastic oligodendroglioma; lung carcinoma tissue;
KW soft tissue leiomyosarcoma; ovary tumour tissue; germ cell tumour tissue;
KW transgenic animal; human; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..90
FT CDS /tag= a
FT CDS 91..3253 /tag= b
FT 3'UTR /product= "Human kinase protein"
FT 2191..3253 /tag= c
XX
US2003175927-A1.
XX
18-SEP-2003.
XX
31-JUL-2002; 2002US-00207973.
XX
16-JAN-2001; 2001US-00759359.
XX
(APPL-) APPLERA CORP.
XX
PI Abu-Threideh J, Gong F, Ketchum KA, Di Francesco V, Beasley EM;
XX
WPI; 2003-898544/82.
XX
DR P-PSDB; ADJ37689.
XX
PT New peptides related to kinase protein subfamily useful for treating disorders associated with abnormal expression of kinase protein in

PT testis, nervous tissue, fetal, lung, ovary tumor tissue.
XX Claim 4; SEQ ID NO 1; 102bp; English.
XX The present invention relates to kinase protein and nucleotides encoding
CC them. The invention is useful in drug screening assays, tissue typing,
CC gene therapy and pharmacogenomic analysis. They are also useful in
CC treating disorders associated with the absence of, inappropriate, or
CC unwanted expression of kinase protein in testis, nervous tissue, fetal,
CC lung, brain anaplastic oligodendroglioma, lung carcinoma tissue, soft
CC tissue leiomyosarcoma, ovary tumour tissue, or germ cell tumour tissue.
CC The invention are useful as models for the development of human
CC therapeutic targets, aid in the identification of therapeutic proteins
CC and serve as targets for the development of human therapeutic agents that
CC modulate protease activity in cells and tissues that express the kinase
CC peptide. The invention is also useful in producing a kinase protein or
CC peptide, and non-human transgenic animals. The present sequence is human
XX kinase cDNA.

XX
SQ Sequence 3253 BP; 991 A; 693 C; 723 G; 846 T; 0 U; 0 Other;
Query Match 100.0%; Score 3253; DB 10; Length 3253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TCGGCGGAGCGAGTGGAGCTCGAGCCAGCTCGTCTCGGCGCGCGCTCGCGTGGCGA 60
DB 1 TCGGCGGAGCGAGTGGAGCTCGAGCCAGCTCGTCTCGGCGCGCGCTCGCGTGGCGA 60

QY 61 AGCCCCCGCGCGCTTCGCGCGCGCTCGGAGTGGCTCCCGGAAAGTGTGCGCCATTGAG 120
DB 61 AGCCCCCGCGCGCTTCGCGCGCGCTCGGAGTGGCTCCCGGAAAGTGTGCGCCATTGAG 120

QY 121 GCCCGAAGCGGAGCGCGAAGAGAGAAACATCCGAAAGAGCGCGAGCTCAACAGAA 180
DB 121 GCCCGAAGCGGAGCGCGAAGAGAGAAACATCCGAAAGAGCGCGAGCTCAACAGAA 180

QY 181 GCTCCCTTAGTCTCTCTCCAGCGCGCACACACACACCGCCACCTTTGGCAGAC 240
DB 181 GCTCCCTTAGTCTCTCTCCAGCGCGCACACACACACCGCCACCTTTGGCAGAC 240

QY 241 CCCACACCCCGGAGCGAGAGGAGATCTGGGATCAGATGATGAGAGCAAGAGGAC 300
DB 241 CCCACACCCCGGAGCGAGAGGAGATCTGGGATCAGATGATGAGAGCAAGAGGAC 300

QY 301 CTTGGGAGCTACTGCAAGTGTGATATCATCCAGTGAATTTGGAGACCTCTTCAATGGC 360
DB 301 CTTGGGAGCTACTGCAAGTGTGATATCATCCAGTGAATTTGGAGACCTCTTCAATGGC 360

QY 361 CGGTATCATGTTATTAGAAAGTGGATGATGATGATGATGATGATGATGATGATGATG 420
DB 361 CGGTATCATGTTATTAGAAAGTGGATGATGATGATGATGATGATGATGATGATGATG 420

QY 421 GATATGAGGGGAAAGATTGTTGCAATGAAGTTGTTAAAGATGCGCCAGCATTTATACG 480
DB 421 GATATGAGGGGAAAGATTGTTGCAATGAAGTTGTTAAAGATGCGCCAGCATTTATACG 480

QY 481 GAGACAGCTTGGATGAATAAATTTGCTCAAAATGTTGTCGAGAAAGTGTATCCAGTGAC 540
DB 481 GAGACAGCTTGGATGAATAAATTTGCTCAAAATGTTGTCGAGAAAGTGTATCCAGTGAC 540

QY 541 CCAACCAAGAGCATGGTGGTCCAGCTCATTTGACGACTTCAAGATTTTCAGGCATGAATGGG 600
DB 541 CCAACCAAGAGCATGGTGGTCCAGCTCATTTGACGACTTCAAGATTTTCAGGCATGAATGGG 600

QY 601 ATACATGTCGATGGTCTTCGAGTACTTGGCCACCATCTCTCAAGTGGATCATCAA 660
DB 601 ATACATGTCGATGGTCTTCGAGTACTTGGCCACCATCTCTCTCAAGTGGATCATCAA 660

QY 661 TCCAACTATCAAGGCTCCCAAGTACGTTGTGGAAGAGATATCAATTCGACAGGCTCTTCAA 720
DB 661 TCCAACTATCAAGGCTCCCAAGTACGTTGTGGAAGAGATATCAATTCGACAGGCTCTTCAA 720

QY 721 GGCTTAGATTACTTACACAGTAAGTCAAGATCATTTATCTACTACATATAAGCCGGAAT 780
DB 721 GGCTTAGATTACTTACACAGTAAGTCAAGATCATTTATCTACTACATATAAGCCGGAAT 780

QY 781 ATCTTGATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 ATCTTGATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840

QY 841 CAGAAAGCAGTCT 900
DB 841 CAGAAAGCAGTCT 900

QY 901 CCTATAGGAAAAATATCTAAAAACAAAAAAGAAAACTGAAAAAAGAAAAAGAGAGCGAG 960
DB 901 CCTATAGGAAAAATATCTAAAAACAAAAAAGAAAACTGAAAAAAGAAAAAGAGAGCGAG 960

QY 961 GCTGAGTATTGGAGAGCGCTTCGAGGAGATGAGAAATTTGGAGCGAGAGCTGAAAGG 1020
DB 961 GCTGAGTATTGGAGAGCGCTTCGAGGAGATGAGAAATTTGGAGCGAGAGCTGAAAGG 1020

QY 1021 AATAATAGAGAAAAATCATCCTCTCAGCTGCACTTCCCAATCACCAGGATGCGAATATC 1080
DB 1021 AATAATAGAGAAAAATCATCCTCTCAGCTGCACTTCCCAATCACCAGGATGCGAATATC 1080

QY 1081 TGCCAGAGGTGAAATCTAAAAACAAACAGGATTAGAGAGCGCGCTGAGGAGAGACTGCA 1140
DB 1081 TGCCAGAGGTGAAATCTAAAAACAAACAGGATTAGAGAGCGCGCTGAGGAGAGACTGCA 1140

QY 1141 AAGGACAACTGGTGAAGCTGAGGACCCAGAGAGAGAAAGAAAGTCTGAGAAAGAAAAACAT 1200
DB 1141 AAGGACAACTGGTGAAGCTGAGGACCCAGAGAGAGAAAGAAAGTCTGAGAAAGAAAAACAT 1200

QY 1201 GAAAAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB 1201 GAAAAAGATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260

QY 1261 TCACCTAAACCAATGCGCATTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1320
DB 1261 TCACCTAAACCAATGCGCATTTGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1320

QY 1321 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380

QY 1381 AATGAGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
DB 1381 AATGAGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440

QY 1441 AATGAGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 1441 AATGAGAAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500

QY 1501 GGATCTTTAGAACCTGTGGCTGCGGCTCTGTGCTCTGTGCTCTGTGCTCTGTGCTCTGT 1560
DB 1501 GGATCTTTAGAACCTGTGGCTGCGGCTCTGTGCTCTGTGCTCTGTGCTCTGTGCTCTGT 1560

QY 1561 CAAAGGAGAGCAGTCTCATTCCTCCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
DB 1561 CAAAGGAGAGCAGTCTCATTCCTCCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620

QY 1621 GATTTGCGAAAGCAAAAAACCCGGGCGAGCTGATCTTTGTTGGTGAATCCCTCGATCCGG 1680
DB 1621 GATTTGCGAAAGCAAAAAACCCGGGCGAGCTGATCTTTGTTGGTGAATCCCTCGATCCGG 1680

QY 1681 AATGAGAGTAAATTTAGAGTAAAAATTTGCTGACCTGGGAAATGCTTTGTTGGTGAATA 1740
DB 1681 AATGAGAGTAAATTTAGAGTAAAAATTTGCTGACCTGGGAAATGCTTTGTTGGTGAATA 1740

QY 1741 CACTTTCAGGAGAGATCCAGAGCGCTCAGTACCCCTCCATAGAGGTTTTTAATAGAGCG 1800
DB 1741 CACTTTCAGGAGAGATCCAGAGCGCTCAGTACCCCTCCATAGAGGTTTTTAATAGAGCG 1800

QY 1801 GGGTACAGCACCCCTCGGAGCATCTGGAGACAGCGGCTGTATGCGAATTTGAGCTGGCAGG 1860

Db 1801 GGGTACAGCACCCCTGGGACATCTGGAGCAGCGGTGTATGGCATTTGAGCTGGCAAG 1860
QY 1861 GGAGATTATTTGTTGAAACACATTTCTGGGGAAGACTATTTCCAGAGACGAAGACACATA 1920
Db 1861 GGAGATTATTTGTTGAAACACATTTCTGGGGAAGACTATTTCCAGAGACGAAGACACATA 1920
QY 1921 GCCCAGCATATGAGCTGTAGGAGTATTTCCAGGCACTTTGCTCTATCTGGAATAAT 1980
Db 1921 GCCCAGCATATGAGCTGTAGGAGTATTTCCAGGCACTTTGCTCTATCTGGAATAAT 1980
QY 1981 TCTCGGGAATTTCTTCAATCGCAGAGAGAACTGCGACACATCAACCAAGCTGAAGCCCTGG 2040
Db 1981 TCTCGGGAATTTCTTCAATCGCAGAGAGAACTGCGACACATCAACCAAGCTGAAGCCCTGG 2040
QY 2041 AGCCTCTTTGATGTAATCTGTGGAAGAGTATGGCTGGCCCAATGAAGATCTGACAGTTT 2100
Db 2041 AGCCTCTTTGATGTAATCTGTGGAAGAGTATGGCTGGCCCAATGAAGATCTGACAGTTT 2100
QY 2101 ACAGATTTTCTGATCCCGATGTTAGAAATGTTTCCAGAAAGAGAGCCCTCAGTGGCGAA 2160
Db 2101 ACAGATTTTCTGATCCCGATGTTAGAAATGTTTCCAGAAAGAGAGCCCTCAGTGGCGAA 2160
QY 2161 TGCCTTCGGCATCTTGGTGTGAATCTTAGCAAAATCTACCAATATTGCAATCTGAGCTA 2220
Db 2161 TGCCTTCGGCATCTTGGTGTGAATCTTAGCAAAATCTACCAATATTGCAATCTGAGCTA 2220
QY 2221 GCAATGTTTCCAGTACATGGACCTTAAACGGTGACTCTCATTTTAAACAGATTACAA 2280
Db 2221 GCAATGTTTCCAGTACATGGACCTTAAACGGTGACTCTCATTTTAAACAGATTACAA 2280
QY 2281 GTGAGCTGGCTTCATCTCAGACCTTATTTGCTTTGAGGTACTGTTGTTGACATTTT 2340
Db 2281 GTGAGCTGGCTTCATCTCAGACCTTATTTGCTTTGAGGTACTGTTGTTGACATTTT 2340
QY 2341 GCTTTTGTGCACTGTGATCTCTGGGAAGGTAGTCTTTTGTCTTCAGCTAAGTAGTTTA 2400
Db 2341 GCTTTTGTGCACTGTGATCTCTGGGAAGGTAGTCTTTTGTCTTCAGCTAAGTAGTTTA 2400
QY 2401 CTGACCATTTTCTTCTGGAACCAATTAACATGTCTCTAAGCATTTGTTCTTGTGTGTG 2460
Db 2401 CTGACCATTTTCTTCTGGAACCAATTAACATGTCTCTAAGCATTTGTTCTTGTGTGTG 2460
QY 2461 ACATTCAAATGTCAATTTTGTGAAGAAATACATTTCCCTTTGTTGTTGCGAGTTT 2520
Db 2461 ACATTCAAATGTCAATTTTGTGAAGAAATACATTTCCCTTTGTTGTTGCGAGTTT 2520
QY 2521 TGTAACTATTTATGAAGAAATATTTTGTAGCTGAGTACTATATTAATTTTACAACTTAAAGAA 2580
Db 2521 TGTAACTATTTATGAAGAAATATTTTGTAGCTGAGTACTATATTAATTTTACAACTTAAAGAA 2580
QY 2581 TTATCAAGTTGGGAACCAAGAAATAGCAAGGAAATGTACAATTTTATCTTCTGCAAA 2640
Db 2581 TTATCAAGTTGGGAACCAAGAAATAGCAAGGAAATGTACAATTTTATCTTCTGCAAA 2640
QY 2641 GGGACATCATCTCTGTATTTATAGTGTATGTAATGACCCCTGTAAATGTACTTTGGATT 2700
Db 2641 GGGACATCATCTCTGTATTTATAGTGTATGTAATGACCCCTGTAAATGTACTTTGGATT 2700
QY 2701 AAATATGGGAGGGGGGACCTCAAAATTTTCAAGAAAGCTTAAAGGAAAGGAAAGTAAGG 2760
Db 2701 AAATATGGGAGGGGGGACCTCAAAATTTTCAAGAAAGCTTAAAGGAAAGGAAAGTAAGG 2760
QY 2761 AAAAATACCTTTATATTAATAATACCCCTTTCTTTGTTTTTTTGTGTTTTTCCATTTTCATATT 2820
Db 2761 AAAAATACCTTTATATTAATAATACCCCTTTCTTTGTTTTTTTGTGTTTTTCCATTTTCATATT 2820
QY 2821 ATTAATACATCTTAAAGCTTGGGAAGCACTATGAAGAAATTAATACATGAAGGATCA 2880
Db 2821 ATTAATACATCTTAAAGCTTGGGAAGCACTATGAAGAAATTAATACATGAAGGATCA 2880
QY 2881 AAAATCATAAATCAAAACCCCTATAGTCCAGCAATTCATTTCTCGCGGTCAACTT 2940

Db 2881 AAAATCATAAATCAAAACCCCTATAGTCCAACGACAATTCATTTCTCGCGGTCAACTT 2940
QY 2941 TTTAACAATCTTATATCTAGTACCTGAGACTCTGGTGTCTAAATATTAATTTCTAAATCTAC 3000
Db 2941 TTTAACAATCTTATATCTAGTACCTGAGACTCTGGTGTCTAAATATTAATTTCTAAATCTAC 3000
QY 3001 CACCAAGTTAGCCCGGTAAATGCTCTCTCGTGAATCTGTCAATACATAAATTTTCT 3060
Db 3001 CACCAAGTTAGCCCGGTAAATGCTCTCTCGTGAATCTGTCAATACATAAATTTTCT 3060
QY 3061 ATTTATTTAGTGGTCTCGTTTATCTTTGCGCCACATCTTTGTTTCACTATTTTCTAGTTA 3120
Db 3061 ATTTATTTAGTGGTCTCGTTTATCTTTGCGCCACATCTTTGTTTCACTATTTTCTAGTTA 3120
QY 3121 CTCTTATCTTTGGGTGAATTAATCTCTCATTAATCTCATATAAATTTCTGAATTTTCT 3180
Db 3121 CTCTTATCTTTGGGTGAATTAATCTCTCATTAATCTCATATAAATTTTCTGAATTTTCT 3180
QY 3181 ACATAAACTACTAGAGCTACCTCACCATCTCTGTTTTTAAACGCGAGAGTTACTATATA 3240
Db 3181 ACATAAACTACTAGAGCTACCTCACCATCTCTGTTTTTAAACGCGAGAGTTACTATATA 3240
QY 3241 ATTACTATTATAA 3253
Db 3241 ATTACTATTATAA 3253
RESULT 2
ABS52846
ID ABS52846 standard; cDNA; 3253 BP.
XX
AC ABS52846;
XX
DT 20-NOV-2002 (first entry)
XX
DE cDNA encoding human SR protein-specific kinase 2, SRPK2.
XX
KW Human; kinase; SRPK2; SR protein-specific kinase 2; gene; ss.
XX
OS Homo sapiens.
XX
EH Key Location/Qualifiers
FT 5'UTR 1..90
FT CDS /*tag= a
FT /*tag= b
FT 3'UTR /product= "SR protein-specific kinase 2 (SRPK2) kinase"
FT 2190..3253
FT /*tag= c
XX
PN US2002094560-A1.
XX
PD 18-JUL-2002.
XX
PF 16-JAN-2001; 2001US-00759359.
XX
PR 16-JAN-2001; 2001US-00759359.
XX
PA (ABUT/) ABU-THREIDEH J.
PA (GONG/) GONG F.
PA (KETCHUM K A.) KETCHUM K A.
PA (DI FRANCESCO V.) DI FRANCESCO V.
PA (BEASLEY E M.) BEASLEY E M.
XX
PI Abu-Threideh J, Gong F, Ketchum KA, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-681805/73.
XX
DR P-PSDB; ABG32447.
XX
PT New SR protein-specific kinase 2 peptides and nucleic acid sequences,
PT useful as models for developing human therapeutic targets, in identifying
PT therapeutic proteins, and in identifying agents that modulate kinase
PT activity.

XX
PS
XX
CC
XX
The invention relates to an isolated human SR protein-specific kinase 2 (SRPK2) peptide. The SR protein-specific kinase 2 peptide and nucleic acid sequences can be used as models for the development of human therapeutic targets, aid in the identification of therapeutic proteins, and as targets for the development of human therapeutic agents that modulate kinase activity in cells and tissues that express the kinase. These may further be used as query sequences to perform a search against sequence databases to identify other family members or related sequences. The peptides can also be used to raise antibodies or to elicit another immune response, as markers for tissues in which the corresponding protein is preferentially expressed, to identify inhibitors of the binding partner/ligand to develop a system to identify inhibitors of the binding interaction, and in pharmacogenomic analysis. The nucleic acids are useful as probes or primers, for expressing antigenic portions of the proteins, for constructing vectors, host cells or transgenic animals expressing the nucleic acids and peptide, for monitoring the effectiveness of modulating compounds on the expression or activity of the kinase gene in clinical trials or in treatment regimen, and as antisense constructs to control kinase gene expression. The present sequence represents the coding sequence of human SRPK2

XX
SQ
Sequence 3253 BP; 991 A; 695 C; 721 G; 846 T; 0 U; 0 Other;

Query Match 99.9%; Score 3249.8; DB 6; Length 3253;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3251; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TCGCGGAGCGAGTGAGGCTGAGCCAGCTCGTCTGGCGCCCGCTCGCGTGGGA 60
DB 1 TCGCGGAGCGAGTGAGGCTGAGCCAGCTCGTCTGGCGCCCGCTCGCGTGGGA 60

QY 61 AGCCCCCCCCCGCTTCCGCGCGCTCGGAATGAGTCCCGGAAGTCTGSCCATTCAG 120
DB 61 AGCCCCCCCCCGCTTCCGCGCGCTCGGAATGAGTCCCGGAAGTCTGSCCATTCAG 120

QY 121 GCCCGAAAGCGGAGCGGCGGAAAGAGAGAAACATCCGAAAGCCGAGGCTCAACAGAAA 180
DB 121 GCCCGAAAGCGGAGCGGCGGAAAGAGAGAAACATCCGAAAGCCGAGGCTCAACAGAAA 180

QY 181 GCTCCTTTTATGCTCTCTCTCCACCGCACACACACACACACACACCTTTGCCAGAC 240
DB 181 GCTCCTTTTATGCTCTCTCTCCACCGCACACACACACACACACCTTTGCCAGAC 240

QY 241 CCCACACCCCGGAGCCAGAGGAGAGATCCTGGATCAGATGATGAGGAGCAAGAGGAC 300
DB 241 CCCACACCCCGGAGCCAGAGGAGAGATCCTGGATCAGATGATGAGGAGCAAGAGGAC 300

QY 301 CTTGCGGACTACTGCAAGGTGGATATCATCCAGTGAAATTTGGAGACCTCTTCAATGGC 360
DB 301 CTTGCGGACTACTGCAAGGTGGATATCATCCAGTGAAATTTGGAGACCTCTTCAATGGC 360

QY 361 CGGTATCATGTTATAGAAAGCTTGGATGGGGGCACTTCTACTGTCTGCTGTGCTGG 420
DB 361 CGGTATCATGTTATAGAAAGCTTGGATGGGGGCACTTCTACTGTCTGCTGTGCTGG 420

QY 421 GATATGCGGGGAAAGATTTGTTGCAATGAAAGTTGTAAAGTGCACAGCATTTATAGC 480
DB 421 GATATGCGGGGAAAGATTTGTTGCAATGAAAGTTGTAAAGTGCACAGCATTTATAGC 480

QY 481 GAGACAGCCTTGGATGAAATTTGCTCAAAATGTTGCGAAGAGTATCCAGTGAC 540
DB 481 GAGACAGCCTTGGATGAAATTTGCTCAAAATGTTGCGAAGAGTATCCAGTGAC 540

QY 541 CCAACCAAGACATGGTGGTCCAGCTCATTCAGATTCAGATTTTCAGSCATGAATGG 600
DB 541 CCAACCAAGACATGGTGGTCCAGCTCATTCAGATTCAGATTTTCAGSCATGAATGG 600

QY 601 ATACATCTCTGCATGGTCTTTCGAAGTACTTGGGCCACCATCTCTCAAGTGGATCATCAA 660
DB 601 ATACATCTCTGCATGGTCTTTCGAAGTACTTGGGCCACCATCTCTCAAGTGGATCATCAA 660

QY 661 TCCAACATATCAAGGCTCCAGTACGTGTGTGAAGAGTATCATTCGACAGGTCCTTCAA 720
DB 661 TCCAACATATCAAGGCTCCAGTACGTGTGTGAAGAGTATCATTCGACAGGTCCTTCAA 720

QY 721 GGGTTAGATTACTTACACAGTAACTGACATCATTCATGACATTAAGACCGGAAAT 780
DB 721 GGGTTAGATTACTTACACAGTAACTGACATCATTCATGACATTAAGACCGGAAAT 780

QY 781 ATCTTGATGTGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840
DB 781 ATCTTGATGTGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 840

QY 841 CAGAAACAGGTGCT 900
DB 841 CAGAAACAGGTGCT 900

QY 901 CCTATAGGAAATATCTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 960
DB 901 CCTATAGGAAATATCTTAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 960

QY 961 GCTCAGTTATTGGAGAGCGCTGAGGAGATGAGAAATTTGGAGCGAGAGCTGAAAGG 1020
DB 961 GCTCAGTTATTGGAGAGCGCTGAGGAGATGAGAAATTTGGAGCGAGAGCTGAAAGG 1020

QY 1021 AAAATAATAGAGAAACATCATCTCAGCTGCACTTCCAAATGACAGGATGCGGATAC 1080
DB 1021 AAAATAATAGAGAAACATCATCTCAGCTGCACTTCCAAATGACAGGATGCGGATAC 1080

QY 1081 TGCCCGAGGTTGAACTTAAACAAACAGGATTAGAGAGCGGCTGAGGCGAGACTGCA 1140
DB 1081 TGCCCGAGGTTGAACTTAAACAAACAGGATTAGAGAGCGGCTGAGGCGAGACTGCA 1140

QY 1141 AAGGACAATGTTGAGCTGAGGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1200
DB 1141 AAGGACAATGTTGAGCTGAGGACCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1200

QY 1201 GAAAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
DB 1201 GAAAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260

QY 1261 TCACCTAAACCAATGGCCATATTGAGAAATGGCCCATTTCTCAGTGGAGCAGCAACTGG 1320
DB 1261 TCACCTAAACCAATGGCCATATTGAGAAATGGCCCATTTCTCAGTGGAGCAGCAACTGG 1320

QY 1321 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380

QY 1381 AATGAG 1440
DB 1381 AATGAG 1440

QY 1441 AATGAG 1500
DB 1441 AATGAG 1500

QY 1501 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560
DB 1501 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1560

QY 1561 CAAG 1620
DB 1561 CAAG 1620

QY 1621 GATTTGCCAAAGCAAAACCCCGGAGCTGATGATGATGATGATGATGATGATGATGATG 1680
DB 1621 GATTTGCCAAAGCAAAACCCCGGAGCTGATGATGATGATGATGATGATGATGATGATG 1680

QY 1681 AATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740
DB 1681 AATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1740

```
1741 CACTTCAGCGAAGACATCCAGACGCGTCAGTACCGCTCCATAGAGGTTTTTAATAGGAGCG 1800
1741 CACTTCAGCGAAGACATCCAGACGCGTCAGTACCGCTCCATAGAGGTTTTTAATAGGAGCG 1800
1801 GGGTACAGCACCCCTCGGACATCTGGAGCAGCGCGTGTATGGCAATTTGAGCTGGCAACG 1860
1801 GGGTACAGCACCCCTCGGACATCTGGAGCAGCGCGTGTATGGCAATTTGAGCTGGCAACG 1860
1861 GGGATATATTTGTTGAAACACATCTCGGGAAGACTATTCGAGAGCAAGACCAATAT 1920
1861 GGGATATATTTGTTGAAACACATCTCGGGAAGACTATTCGAGAGCAAGACCAATAT 1920
1921 GCCCAGATCATAGAGCTGTAGGAGTATTCGAGGAGCACTTTCTCTATCTGGAATAT 1980
1921 GCCCAGATCATAGAGCTGTAGGAGTATTCGAGGAGCACTTTCTCTATCTGGAATAT 1980
1981 TCTCGGGAATTTCTCAATCGCAGAGGAGAACTCGGACACATCACCAAGCTGAAGCCCTCG 2040
1981 TCTCGGGAATTTCTCAATCGCAGAGGAGAACTCGGACACATCACCAAGCTGAAGCCCTCG 2040
2041 AGCCTCTTGATGATCTGTGGAAAGATAGGCTGGCCCGCATGAAGATGTCGACAGTTT 2100
2041 AGCCTCTTGATGATCTGTGGAAAGATAGGCTGGCCCGCATGAAGATGTCGACAGTTT 2100
2101 ACAGATTTCTGATCCGATGTTAGAAATGTTCCAGAAACGAGCCCTCAGCTGCGGAA 2160
2101 ACAGATTTCTGATCCGATGTTAGAAATGTTCCAGAAACGAGCCCTCAGCTGCGGAA 2160
2161 TGCTTTGCGGATCTTTGCTGTTGAATTTCTAGCAATTTCTACCAATATTGCAATCTGAGCTA 2220
2161 TGCTTTGCGGATCTTTGCTGTTGAATTTCTAGCAATTTCTACCAATATTGCAATCTGAGCTA 2220
2221 GCAAAATGTTCCAGTACATTTGGAACCTTAAACGGTGAATCTCATTTCTTAAACAGGATTACAA 2280
2221 GCAAAATGTTCCAGTACATTTGGAACCTTAAACGGTGAATCTCATTTCTTAAACAGGATTACAA 2280
2281 GTGAGCTGGCTTCATCTCAGACCTTTATTTGCTTTGAGGACTGTTGTTGACATTTT 2340
2281 GTGAGCTGGCTTCATCTCAGACCTTTATTTGCTTTGAGGACTGTTGTTGACATTTT 2340
2341 GCTTTTGTGCACTGTGATCTCGGGAAGGAGTGTCTTTGCTTCAGCTAAGTAGTTTA 2400
2341 GCTTTTGTGCACTGTGATCTCGGGAAGGAGTGTCTTTGCTTCAGCTAAGTAGTTTA 2400
2401 CTGACCAATTTCTCTGGAACCAATTAACATGCTCTAAGCATGTTTCTGTGTTGTG 2460
2401 CTGACCAATTTCTCTGGAACCAATTAACATGCTCTAAGCATGTTTCTGTGTTGTG 2460
2461 ACATTCAGATGTCATTTTGTGAATGAAATACTTTCCCTTTGTTGTTGCGAGTTT 2520
2461 ACATTCAGATGTCATTTTGTGAATGAAATACTTTCCCTTTGTTGTTGCGAGTTT 2520
2521 TGTAACATTTTATGAAGAAATTTTGTAGCTGAGTACTATATATTTTACAACTTTAAGAAA 2580
2521 TGTAACATTTTATGAAGAAATTTTGTAGCTGAGTACTATATATTTTACAACTTTAAGAAA 2580
2581 TTATCAAGTTGGGAACCAAGAAATAGCAAGGAAATGTACAAATTTTATCTTCGSCAAA 2640
2581 TTATCAAGTTGGGAACCAAGAAATAGCAAGGAAATGTACAAATTTTATCTTCGSCAAA 2640
2641 GGGACATCAATCTCTGTTATATAGTGTATGTAATGACCCCTGTAAATGTTTGTGATTT 2700
2641 GGGACATCAATCTCTGTTATATAGTGTATGTAATGACCCCTGTAAATGTTTGTGATTT 2700
2701 AAATATGGAGGGGGGACTCAAAATTTGAGAAAGCTTAAAGAAAGGAAAGTAAAGG 2760
2701 AAATATGGAGGGGGGACTCAAAATTTGAGAAAGCTTAAAGAAAGGAAAGTAAAGG 2760
2761 AAAAATFACTTTATATTAATAACCTTTCTTTGTTTTTTGTTTTTCTTATTTTCAATAT 2820
2761 AAAAATFACTTTATATTAATAACCTTTCTTTGTTTTTTGTTTTTCTTATTTTCAATAT 2820
2821 ATTAAATACACTTAACTGTCGGAAGACATATGAAAAAATTAATACATGAAAGGATCA 2880
```

```
Db 2821 ATTAAATACACTTAACTGTCGGAAGCACTATGAAAAAATTAATACATGAAAGGATCA 2880
QY 2881 AAAATCATAAATCAAAACCCCACTATAGTCCAACGACAAATTCATTCTCGCGGTCAACTT 2940
Db 2881 AAAATCATAAATCAAAACCCCACTATAGTCCAACGACAAATTCATTCTCGCGGTCAACTT 2940
QY 2941 TTTAAACATCTTATATCTAGTACCTGAGACTCTGGTGTCTCAATATATTAATTTCTAAATCTAC 3000
Db 2941 TTTAAACATCTTATATCTAGTACCTGAGACTCTGGTGTCTCAATATTAATTTCTAAATCTAC 3000
QY 3001 CACCAAGTTAGGCCGCTAATGCTCTCTCGTGAATCTGTCATACATAAATTTCTGAAATTTTCT 3060
Db 3001 CACCAAGTTAGGCCGCTAATGCTCTCTCGTGAATCTGTCATACATAAATTTCTGAAATTTTCT 3060
QY 3061 ATTTATTTAGTGGGTCTCGTTTATCTTTTCCGCCACATCTTTGTTCACTATTTTCTAGTTA 3120
Db 3061 ATTTATTTAGTGGGTCTCGTTTATCTTTTCCGCCACATCTTTGTTCACTATTTTCTAGTTA 3120
QY 3121 CTCTTATCTTTGGGCTGATTAATCCTTCTCATATATCTCATATAAATTTCTGAAATTTTCT 3180
Db 3121 CTCTTATCTTTGGGCTGATTAATCCTTCTCATATATCTCATATAAATTTCTGAAATTTTCT 3180
QY 3181 ACATAAACTACTAGAGTACCTCACCATCTCTGTTTTTAAACGCGAGCAGTTTACTATATA 3240
Db 3181 ACATAAACTACTAGAGTACCTCACCATCTCTGTTTTTAAACGCGAGCAGTTTACTATATA 3240
QY 3241 ATTACTATTATAA 3253
Db 3241 ATTACTATTATAA 3253

RESULT 3
AAx89851
ID: AAX89851 standard; cDNA; 2791 Bp.
XX
AC AAX89851;
XX
DT 08-OCT-1999 (first entry)
XX
DE Human protein kinase (HPKM)-2 encoding nucleotide (clone ID 307624).
XX
KW Human protein kinase molecule; HPKM; human; protein kinase;
KW phosphate group; cancer; immune disorder; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 173..2239
FT FT /*tag= a
FT FT /product= "HPKM-2"
XX
PN WO938981-A2.
XX
PD 05-AUG-1999.
XX
PF 12-JAN-1999; 99WO-US0000661.
XX
PR 30-JAN-1998; 98US-00016000.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Bandman O, Hillman JL, Lal P, Akerblom IE, Shah P, Corley NC;
PI Guegler KJ;
XX
PI WPI: 1999-479190/40.
DR P-PSDB; AAY27053.
XX
PT New human protein kinase molecules useful for treating or preventing
PT cancer or an immune disorder.
XX
PS Claim 8; Page 73-74; 77pp; English.
XX
```


Qy	1931	TAGAGCTGCTAGG	CAGTATTCCAA	GGCACCTTTGCTCTATCTGGA	AAATATTTCGGGAAT	1990
Db	1980	TAGAGCTGCTAGG	CAGTATTCCAA	GGCACCTTTGCTCTATCTGGA	AAATATTTCGGGAAT	2039
Qy	1991	TCCTTCAATCGCAG	AGGAGAACTCGG	CACACATCACCAAGCTG	AGCCCTCGAGCCTCTTTG	2050
Db	2040	TCCTTCAATCGCAG	AGGAGAACTCGG	CACACATCACCAAGCTG	AGCCCTCGAGCCTCTTTG	2099
Qy	2051	ATGPACTTGTG	AAAAAGTATGGCTGG	CCCCCATGAAGATGCTGCA	CAGTTTACAGATTTC	2110
Db	2100	ATGTACTTGTG	AAAAAGTATGGCTGG	CCCCCATGAAGATGCTGCA	CAGTTTACAGATTTC	2159
Qy	2111	TGATCCGAGTGT	TAGAAATGGTTC	CCAGAAAACGAGCTCAGCTGG	GGAATGCTTCGGC	2170
Db	2160	TGATCCGAGTGT	TAGAAATGGTTC	CCAGAAAACGAGCTCAGCTGG	GGAATGCTTCGGC	2219
Qy	2171	ATCCTTGGT	TGAAATCTTTAG	CAAAATCTTACCAATATTGCA	ATTTCTGAGCTAGCAAAATGTC	2230
Db	2220	ATCCTTGGT	TGAAATCTTTAG	CAAAATCTTACCAATATTGCA	ATTTCTGAGCTAGCAAAATGTC	2279
Qy	2231	CCAGTACATTGG	GAACCTTAAACGGT	GACTCTCATCTTTTAAACAGGA	ATTACAAGTGAAGCTGGC	2290
Db	2280	CCAGTACATTGG	GAACCTTAAACGGT	GACTCTCATCTTTTAAACAGGA	ATTACAAGTGAAGCTGGC	2339
Qy	2291	TTCAATCCTCAG	ACCTTTATTTTGGCTTTGAG	TACTGTTGTTGACATTTTGGCTTTTGTG		2350
Db	2340	TTCAATCCTCAG	ACCTTTATTTTGGCTTTGAG	TACTGTTGTTGACATTTTGGCTTTTGTG		2399
Qy	2351	CACGTGTGATC	CTGGGGAAGGTAGTCTTTTGTCTT	CAGCTAAAGTAGTTTACTGACCAATTT		2410
Db	2400	CACGTGTGATC	CTGGGGAAGGTAGTCTTTTGTCTT	CAGCTAAAGTAGTTTACTGACCAATTT		2459
Qy	2411	TCCTTCTGGAA	ACAATAACATGCTCTAAG	CAATGTTTCTTGTTGTGTGACATTTCAAAT		2470
Db	2460	TCCTTCTGGAA	ACAATAACATGCTCTAAG	CAATGTTTCTTGTTGTGTGACATTTCAAAT		2519
Qy	2471	GTCAATTTTTT	TGAAATGAAAAATAC	TTTCCCCTTTGTGTTTGGCAGGTTTGTAACTATT		2530
Db	2520	GTCAATTTTTT	TGAAATGAAAAATAC	TTTCCCCTTTGTGTTTGGCAGGTTTGTAACTATT		2579
Qy	2531	TATGAAGAAAT	ATTTTTAGCTGAGTACTATATAATTT	TACAATCTTTAAAGAAATTTATCAAGTT		2590
Db	2580	TATGAAGAAAT	ATTTTTAGCTGAGTACTATATAATTT	TACAATCTTTAAAGAAATTTATCAAGTT		2639
Qy	2591	GGGAAACCA	AGNAATAGCAAGG	AAATGTACAATTTTTATCTTCTGGCA	AAAGGACATCAT	2650
Db	2640	GGAAACCA	AGNAATAGCAAGG	AAATGTACAATTTTTATCTTCTGGCA	AAAGGACATCAT	2696
Qy	2651	TCCTGTATTT	ATAGTGTATGTAATGCA	CCCTGTAAATGTTACTTTGATT	TAATAATGGGA	2710
Db	2697	TCCTGTATTT	ATAGTGTATGTAATGCA	CCCTGTAAATGTTACTTTGATT	TAATAATGGGA	2756
Qy	2711	GGGGGGACT	CAAATTTTCAGAAA	AGCTAAAA	AAAAA	2746
Db	2757	-GGGGGGACT	CAAATTTTCAGAAA	AGCTAAAA	AAAAA	2791

RESULT 4

RESULI 4
AAI.54214

AAL54214
 ID AAL54214 standard: DNA: 3702 BP.

XX
XX
4174CTMY DT

AC AAL54214:

XX
XXXXXX
1470689

27-MAR-2003 (first entry)

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

DE SR protein-specific kinase

XX

KW Cytostatic; p53 pathway-mo

KW lung cancer; breast cancer

KW colon cancer; ovarian cancer

XX

OS Homo sapiens.

XX

FN	WO200299427-A1.
XX	
PD	12-DEC-2002.
XX	
Pf	03-JUN-2002; 2002WO-US017525.
XX	
PR	05-JUN-2001; 2001US-0296076P.
PR	10-OCT-2001; 2001US-0328605P.
PR	15-FEB-2002; 2002US-0357253P.
XX	(EXEL-) EXELIXIS INC.
PA	
XX	Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI	WPI, 2003-156865/15.
XX	
DR	Identifying candidate p53 pathway-modulating agents useful as therapeutic
XX	targets for disorders related with defective p53 function, by screening
PT	for agents modulating serine/arginine protein-specific kinase activity.
PT	
PS	Disclosure; Page 109-111; 137pp; English.
XX	
CC	The invention relates to a novel method for identifying candidate p53
CC	pathway-modulating agents. The novel method comprises screening for
CC	agents that modulate the activity of a serine/arginine protein-specific
CC	kinase. The invention provides methods for utilizing p53 modifier genes
CC	and polypeptides to identify candidate therapeutic agents that can be
CC	used in the treatment of disorders associated with defective p53
CC	function. The methods are also useful for modulating a p53 pathway in a
CC	mammalian cell, or for diagnosing or treating a disease associated with
CC	defective p53 function, e.g. cancers such as breast cancer, colon cancer,
CC	lung cancer or ovarian cancer. The serine/arginine protein-specific
CC	kinase (SRPK) polypeptides and nucleic acids are useful for identifying
CC	and testing agents that modulate SRPK function. The animal models are
CC	useful for in vivo assays to test the activity of a candidate p53-
CC	modulating agent, or to assess the role of SRPK in a p53 pathway process
CC	such as apoptosis or cell proliferation. This polynucleotide sequence
CC	represents the DNA of a serine/arginine-rich domain protein-specific
CC	kinase-1 of the invention
XX	
SQ	Sequence 3702 BP; 1130 A; 716 C; 823 G; 1033 T; 0 U; 0 Other;
	Query Match 77.9%; Score 2533.8; DB 8; Length 3702;
	Best Local Similarity 99.2%; Pred. No. 0;
	Matches 2568; Conservative 0; Mismatches 17; Indels 4; Gaps 2;
Qy	153 TCGAATAAACGCCGAGCCTCAACAGAAGCTCCTTTAGTTCCTCTCTCCACGGCCACC 212
Dd	202 TTCAGAAGGCCGGAGCCTCAACAGAAGCTCCTTTAGTTCCTCTCTCCACGGCCACC 261
Qy	213 ACCACACACCGGCACCTTTGCCAGACCCACACCCCGGAGCCAGGAGGATCCT 272
Dd	262 ACCACACACCGGCACCTTTGCCAGACCCACACCCCGGAGCCAGGAGGATCCT 321
Qy	273 GGGATCATGATGATGAGGACAAGAGCACCTCGCGACTACTGCAAAGGTGGATATCATCC 332
Dd	322 GGGATCATGATGATGAGGACAAGAGCACCTCGCGACTACTGCAAAGGTGGATATCATCC 381
Qy	333 AGTGAAAAATTGGAGACCTCTTCAATGCCCGTATCATGTTATTAGAAAGCTTGGATGGGG 392
Dd	382 AGTGAAAAATTGGAGACCTCTTCAATGCCCGGTATCATGTTATTAGAAAGCTTGGATGGGG 441
Qy	393 GCATCTCTACTGTCTGTGGTGTGTGGGATATGCGAGGGAAAAGATTTGTTCGAATGAA 452
Dd	442 GCATCTCTACTGTCTGTGGTGTGTGGGATATGCGAGGGAAAAGATTTGTTCGAATGAA 501
Qy	453 AGTTGTAAAAAGTGCCCGACATTTATACGGAGACAGCCCTTGGATGAAATAAAATTCCTCAA 512
Dd	502 AGTTGTAAAAAGTGCCCGACATTTATACGGAGACAGCCCTTGGATGAAATAAAATTCCTCAA 561
Qy	513 ATGTGTTTCGAGAAAGTAGTCCAGTAGTACCCAAACAAAGACATGGTGTGTCAGCTCATTTGA 572
Dd	562 ATGTGTTTCGAGAAAGTAGTCCAGTAGTACCCAAACAAAGACATGGTGTGTCAGCTCATTTGA 621

QY	573	CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCTGATGTCTTCGAGTACTTGG	632	QY	1653	CTTTGTTGGTGAATCCCTGGATCCCGGAATGCGAATAAAATTAGAGTAAAAATTGCTGA	1712
Db	622	CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCTGCTGATGTCTTCGAGTACTTGG	681	Db	1702	CTTTGTTGGTGAATCCCTGGATCCCGGAATGCGAATAAAATTAGAGTAAAAATTGCTGA	1761
QY	633	CAACATCTCTCAAGTGGATCATCAATCAACTATCAAGGCCTCCCAAGTACCTTGTGT	692	QY	1713	CTTGGGAATGCTTGTGGTGCATAAACACTTACGGAAGACATCCAGACGCGTCAAGTA	1772
Db	682	CAACATCTCTCAAGTGGATCATCAATCAACTATCAAGGCCTCCCAAGTACCTTGTGT	741	Db	1762	CTTGGGAATGCTTGTGGTGCATAAACACTTACGGAAGACATCCAGACGCGTCAAGTA	1821
QY	693	GAAAGATGATCAATTCGACAGGTCCTTCAAGGGTTAGATTA CTTA CACAGTAAGTCAAGAT	752	QY	1773	CGCTCCATAGAGGTTTTTAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGAGGAC	1832
Db	742	GAAAGATGATCAATTCGACAGGTCCTTCAAGGGTTAGATTA CTTA CACAGTAAGTCAAGAT	801	Db	1822	CGCTCCATAGAGGTTTTTAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGAGGAC	1881
QY	753	CATTCACTAGCATATAAGCGCGGAAAAATATCTTGATGTGTGATGATGATGATGATG	812	QY	1833	GGCTGTATGGCAATTTGAGCTGCGCAACCGGAGATTTATTTGTTTGAACACATCTTGGGA	1892
Db	802	CATTCACTAGCATATAAGCGCGGAAAAATATCTTGATGTGTGATGATGATGATGATG	861	Db	1882	GGCTGTATGGCAATTTGAGCTGCGCAACCGGAGATTTATTTGTTTGAACACATCTTGGGA	1941
QY	813	AGAAATGGCAGCTGAGGCGACCTGAGTGGCAGAAAGCAGGTCTCTCTCTCTCTCAGGTC	872	QY	1893	AGACTATTCAGAGACGAAAGACACATAGCCACATCATAGAGCTGTAGGAGATTTCC	1952
Db	862	AGAAATGGCAGCTGAGGCGACCTGAGTGGCAGAAAGCAGGTCTCTCTCTCTCTCAGGTC	921	Db	1942	AGACTATTCAGAGACGAAAGACACATAGCCACATCATAGAGCTGTAGGAGATTTCC	2001
QY	873	TGCACTGAGTACGGCTCCACAGCAAAACCTTATAGGAAAAATATCTTAAAAACAAAAAGAA	932	QY	1953	AAGGCACCTTTGCTCTATCTGGAAAAATATCTCGGGAAATCTTCAATCGCAGAGGAACT	2012
Db	922	TGCACTGAGTACGGCTCCACAGCAAAACCTTATAGGAAAAATATCTTAAAAACAAAAAGAA	981	Db	2002	AAGGCACCTTTGCTCTATCTGGAAAAATATCTCGGGAAATCTTCAATCGCAGAGGAACT	2061
QY	933	AAACTGAAAAAGAAACAGAGGCGAGCTGAGTATTTGGAGAGCGCGCTGCGAGGAT	992	QY	2013	GCAGACATCAACCAAGCTGAGCCCTGGAGCCTCTTTGATGTACTTTGTGAAAAAGTATGG	2072
Db	982	AAACTGAAAAAGAAACAGAGGCGAGCTGAGTATTTGGAGAGCGCGCTGCGAGGAT	1041	Db	2062	GCAGACATCAACCAAGCTGAGCCCTGGAGCCTCTTTGATGTACTTTGTGAAAAAGTATGG	2121
QY	993	AGAAAGAAATTTGGACGAGAGCTGAAAGGAAAAATATAGAGAAAAACATCACTCAGCTGC	1052	QY	2073	CTGGCCCCATAGAGATGTGCAACAGTTTACAGATTTCTTGATCCCGATGTTAGAAATGGT	2132
Db	1042	AGAAAGAAATTTGGACGAGAGCTGAAAGGAAAAATATAGAGAAAAACATCACTCAGCTGC	1101	Db	2122	CTGGCCCCATAGAGATGTGCAACAGTTTACAGATTTCTTGATCCCGATGTTAGAAATGGT	2181
QY	1053	ACCTTCCATAGACAGATGGCGAATATCTGCCAGAGGTGAACCTAAAAACACAGGAT	1112	QY	2133	TCCAGAAAAACGAGCCTCAGCTGCGGAAATGCTTCCGCAATGTTCCAGATACATTTGAACTT	2192
Db	1102	ACCTTCCATAGACAGATGGCGAATATCTGCCAGAGGTGAACCTAAAAACACAGGAT	1161	Db	2182	TCCAGAAAAACGAGCCTCAGCTGCGGAAATGCTTCCGCAATGTTCCAGATACATTTGAACTT	2241
QY	1113	AGAGGAGCGGCTGAGGACAGAGCTGCAAGGACAAATGGTGAAGCTGAGGACACAGGAA	1172	QY	2193	AAATTTACCAATATTTGCAATTTCTGAGCTAGCAAAATGTTCCAGATACATTTGAACTT	2252
Db	1162	AGAGGAGCGGCTGAGGACAGAGCTGCAAGGACAAATGGTGAAGCTGAGGACACAGGAA	1221	Db	2242	AAATTTACCAATATTTGCAATTTCTGAGCTAGCAAAATGTTCCAGATACATTTGAACTT	2301
QY	1173	GAAAGAAAGATGCTGAGAAAGAAACATTTGAAAGAAATGAAAGATGATGATGATGATG	1232	QY	2253	TGACTCTCATTTCTTTAAACAGGATTAACAGTGAAGCTGGCTTCACTCCAGACCTTTATTT	2312
Db	1222	GAAAGAAAGATGCTGAGAAAGAAACATTTGAAAGAAATGAAAGATGATGATGATGATG	1281	Db	2302	TGACTCTCATTTCTTTAAACAGGATTAACAGTGAAGCTGGCTTCACTCCAGACCTTTATTT	2361
QY	1233	TGCGAACATAGACCTTACGTTAGTAAATCACTTAAACCAATGGCCATATTCAGAAATGG	1292	QY	2313	GCTTTGAGGTACTGTTGTTGACATTTTGTGTTTGTGCACTGTGATCTCTGGGGAAGGTT	2372
Db	1282	TGCGAACATAGACCTTACGTTAGTAAATCACTTAAACCAATGGCCATATTCAGAAATGG	1341	Db	2362	GCTTTGAGGTACTGTTGTTGACATTTTGTGTTTGTGCACTGTGATCTCTGGGGAAGGTT	2421
QY	1293	CCCATTTCTCACTGGAGCAGCAACTGGACGATGATGATGATGATGATGATGATGATGATG	1352	QY	2373	AGTCTTTTGTCTTCTCAGCTAAGTATTTACTGACCATTTCTTCTGGAACAAATTAACATGT	2432
Db	1342	CCCATTTCTCACTGGAGCAGCAACTGGACGATGATGATGATGATGATGATGATGATGATG	1401	Db	2422	AGTCTTTTGTCTTCTCAGCTAAGTATTTACTGACCATTTCTTCTGGAACAAATTAACATGT	2481
QY	1353	TCCTGAGGAAATATATCTTGATGAGCCAAATGCGAAAGTATTAACATATATAGCAGCTC	1412	QY	2433	CTCTAAGCAATGTTTCTTGTGTTGTTGACATTTCAATGTTCAATTTTGTGAATGAAATAAT	2492
Db	1402	TCCTGAGGAAATATATCTTGATGAGCCAAATGCGAAAGTATTAACATATATAGCAGCTC	1461	Db	2482	CTCTAAGCAATGTTTCTTGTGTTGTTGACATTTCAATGTTCAATTTTGTGAATGAAATAAT	2541
QY	1413	CTATGACAAATTCATGTTGATTTGCCAAATGGACGACATATAAAATTTCCCGAGTCAAGTT	1472	QY	2493	ACCTTTCCCTTTGCTTTTGGCAGGTTTGTAACTTATTTATGAAGAAATAATTTAGCTGA	2552
Db	1462	CTATGACAAATTCATGTTGATTTGCCAAATGGACGACATATAAAATTTCCCGAGTCAAGTT	1521	Db	2542	ACCTTTCCCTTTGCTTTTGGCAGGTTTGTAACTTATTTATGAAGAAATAATTTAGCTGA	2601
QY	1473	CCGAGGTTTTCCACCTCGTTTCTCTGATTCCTTGAATTCCTTGAACCTGTGGCTGCGGCTCTGT	1532	QY	2553	GTACTATATATAATTAATCTTAAAGAAATTAATCAAGTTGGGAAACCAAGAAATAGCAAGG	2612
Db	1522	CCGAGGTTTTCCACCTCGTTTCTCTGATTCCTTGAATTCCTTGAACCTGTGGCTGCGGCTCTGT	1581	Db	2602	GTACTATATATAATTAATCTTAAAGAAATTAATCAAGTTGGGAAACCAAGAAATAGCAAGG	2658
QY	1533	GCTTTCTGAGGATCAACATCTAGCAGCAAGAGGAGGAGTCCATCCCATGACAGAAAG	1592	QY	2613	GAAATGTACAAATTTTATCTTCTGGCAAAAGGAGCATCTTCTCTGTATTTATAGTGTATGTA	2672
Db	1582	GCTTTCTGAGGATCAACATCTAGCAGCAAGAGGAGGAGTCCATCCCATGACAGAAAG	1641	Db	2659	GAAATGTACAAATTTTATCTTCTGGCAAAAGGAGCATCTTCTCTGTATTTATAGTGTATGTA	2718
QY	1593	CAGAACGGTTTTCCAGCTCCAGTATCTGGGATTTTGCAAAAACAAAAACCCGGGAGCTGA	1652	QY	2673	ATGCACCTGTAAATGTTTACTTTTGAATTAATAATATGGAGGGGGAGCTCAATTTTCAAGAA	2732
Db	1642	CAGAACGGTTTTCCAGCTCCAGTATCTGGGATTTTGCAAAAACAAAAACCCGGGAGCTGA	1701	Db	2719	ATGCACCTGTAAATGTTTACTTTTCCATTTAAATATATGGGA - GGGGAGCTCAATTTTCAAGAA	2777
				QY	2733	AGCTAAAAA	2741


```

Db      2778 AGTACCAA 2786
|||||
RESULT 5
ADJ96579
ID      ADJ96579 standard; DNA; 3715 BP.
XX      AC
XX      ADJ96579;
XX      DT
XX      06-MAY-2004 (first entry)
XX      DE
XX      Human SRPK2 DNA SeqID 36.
XX      gene; ds; kinase; human; SNP; single nucleotide polymorphism;
KW      tyrosine protein kinase; serine/threonine protein kinase; PK; STK;
KW      gene therapy; cancer; immune-related disease; cardiovascular disease;
KW      brain; neuronal associated disease; metabolic; inflammatory disorder;
KW      cytostatic; neuroprotective; immunomodulator; antiinflammatory; SRPK2.
XX      OS
XX      Homo sapiens.
XX      24.
XX      FH
XX      Key      Location/Qualifiers
XX      FT      variation      replace(1969,t)
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      /tag= a
XX      FT      variation      replace(1996,g)
XX      FT      variation      /tag= b
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(1999,g)
XX      FT      variation      /tag= c
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2001,a)
XX      FT      variation      /tag= d
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2014,a)
XX      FT      variation      /tag= e
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2016,a)
XX      FT      variation      /tag= f
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2017,c)
XX      FT      variation      /tag= g
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2021,a)
XX      FT      variation      /tag= h
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2023,g)
XX      FT      variation      /tag= i
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2029,g)
XX      FT      variation      /tag= j
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2035,c)
XX      FT      variation      /tag= k
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2040,a)
XX      FT      variation      /tag= l
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2044,g)
XX      FT      variation      /tag= m
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2047,t)
XX      FT      variation      /tag= n
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      FT      variation      replace(2219,t)
XX      FT      variation      /tag= o
XX      FT      variation      /standard_name= "Single nucleotide polymorphism"
XX      PN      WO2004006838-A2.
XX      XX      22-JAN-2004.

```

```

XX      15-JUL-2003; 2003WO-US021730.
XX      PF
XX      PR      15-JUL-2002; 2002US-0395632P.
XX      PA
XX      (SUGB-) SUGEN INC.
XX      PI
XX      Whyte D, Manning G, Caenepeel S;
XX      WPI; 2004-122753/12.
XX      DR      P-PSDB; ADJ96645.
XX      PT
XX      New nucleic acid molecule encoding a kinase polypeptide, useful for
XX      preparing a composition for treating diseases or disorders, e.g., cancer,
XX      or neurological, immunological or inflammatory disorders.
XX      Example 1; SEQ ID NO 36; 365pp; English.
XX      CC      This invention relates to a novel isolated, enriched or purified nucleic
XX      acid molecule that encodes a kinase polypeptide. Specifically, it relates
XX      to human tyrosine and serine/threonine protein kinases (PK's and STK's),
XX      as well as protein kinase-like enzymes. The present invention describes
XX      screening methods to identify agonists, antagonists and antibodies that
XX      can be used to modulate the activity or function of the mammalian kinase
XX      enzymes. As such, these compositions can be used for gene therapy
XX      purposes to treat diseases or disorders including cancer, immune-related
XX      diseases, cardiovascular disease, brain or neuronal associated disease,
XX      metabolic and inflammatory disorders. Accordingly, they exhibit
XX      cytostatic, neuroprotective, immunomodulator and antiinflammatory
XX      activities. This polynucleotide sequence is a human kinase DNA sequence
XX      of the invention.
XX      SQ      Sequence 3715 BP; 1135 A; 719 C; 825 G; 1036 T; 0 U; 0 Other;
XX
XX      Query Match      77.9%; Score 2533.8; DB 12; Length 3715;
XX      Best Local Similarity 99.2%; Pred. No. 0;
XX      Matches 2568; Conservative 0; Mismatches 17; Indels 4; Gaps 2;
XX
XX      QY      153 TCCGAAAAGCCGAGCCTCAACAGAAAGCTCCTTTAGTTCCTCTCTCCACCGCCACC 212
XX      DB      |||||
XX      208 TTCCAGAAAGCCGAGCCTCAACAGAAAGCTCCTTTAGTTCCTCTCTCCACCGCCACC 267
XX      QY      213 ACCACCCACCCGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGATCCT 272
XX      DB      |||||
XX      268 ACCACCCACCCGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAGGAGATCCT 327
XX      QY      273 GGGATCAGATGATGAGGAGCAGCCTCTGGGACTACTGCAAGAGTGGATATCATCC 332
XX      DB      |||||
XX      328 GGGATCAGATGATGAGGAGCAGAGGACCTCTGGGACTACTGCAAGAGTGGATATCATCC 387
XX      QY      333 AGTGAATAATTGAGACCTCTTCAATGGCGGATCATGTTATTAGAAAGCTTGGATGGGG 392
XX      DB      |||||
XX      388 AGTGAATAATTGAGACCTCTTCAATGGCGGATCATGTTATTAGAAAGCTTGGATGGGG 447
XX      QY      393 GCACCTTCTACTGCTGCTGGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 452
XX      DB      |||||
XX      448 GCACCTTCTACTGCTGCTGGGATATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 507
XX      QY      453 AGTGTGTAATAAGTGCACGATTTATACGGAGACAGCCTTTGGATGAATAAATTGCTCAA 512
XX      DB      |||||
XX      508 AGTGTGTAATAAGTGCACGATTTATACGGAGACAGCCTTTGGATGAATAAATTGCTCAA 567
XX      QY      513 ATGTGTTCCGAGAAAGTATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA 572
XX      DB      |||||
XX      568 ATGTGTTCCGAGAAAGTATCCAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA 627
XX      QY      573 GCACCTTCAAGATTTAGGAGATGAATGGGATACATGCTGCTGATGGTCTTGAAGTACTTGG 632
XX      DB      |||||
XX      628 GCACCTTCAAGATTTAGGAGATGAATGGGATACATGCTGCTGATGGTCTTGAAGTACTTGG 687
XX      QY      633 CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCTCCCAAGTACTGTTGTGT 692
XX      DB      |||||
XX      688 CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCTCCCAAGTACTGTTGTGT 747

```

693	GAAGAGTATCAATCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGCAGAT	752	CGCTCCATAGAGGTTTAAATAGAGCGGGGTACAGCAACCCCTCGGACACATCTGGAGCAC	1832
748	GAAGAGTATCAATCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGCAGAT	807	CGCTCCATAGAGGTTTAAATAGAGCGGGGTACAGCAACCCCTCGGACACATCTGGAGCAC	1887
753	CATTTCATCTGACATTAAGCGGAAATATCTTGATGTGTGGATGATGCATATGTGAG	812	GGCGTGTATGSCATTTGAGCTGGCAACGSGAGATTATTTGTTTGAACACACATTTCTGGGGA	1892
808	CATTTCATCTGACATTAAGCGGAAATATCTTGATGTGTGGATGATGCATATGTGAG	867	GGCGTGTATGSCATTTGAGCTGGCAACGSGAGATTATTTGTTTGAACACACATTTCTGGGGA	1947
813	AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAAGTGTCTCTCTTCAGGGTC	872	AGACTATTCAGAGACGAAGACCATAGCCCCACATATAGAGCTGTCTAGGCAGTATTC	1952
868	AAGAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAAGTGTCTCTCTTCAGGGTC	927	AGACTATTCAGAGACGAAGACCATAGCCCCACATATAGAGCTGTCTAGGCAGTATTC	2007
873	TGCAGTGAGTACGGCTCCACAGCAACACCTTATAGGAAATATCTATAAAACAAAAAGAA	932	AAGGCATTTTGTCTATCTGGAATAATTTCTCGGGAATTTCTCAATCGCAGAGGAAC	2067
928	TGCAGTGAGTACGGCTCCACAGCAACACCTTATAGGAAATATCTATAAAACAAAAAGAA	987	AAGGCATTTTGTCTATCTGGAATAATTTCTCGGGAATTTCTCAATCGCAGAGGAAC	2067
933	AAAACTGAAAAAGAAACAGAGAGCGAGCTGAGTTATTTGGAGAGCGCTGCAGAGAT	992	GCAGACATCAACAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGGAATAAGTATGG	2072
988	AAAACTGAAAAAGAAACAGAGAGCGAGCTGAGTTATTTGGAGAGCGCTGCAGAGAT	1047	GCAGACATCAACAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGGAATAAGTATGG	2127
993	AGAAGAAATTTGAGCGGAGAGCTGAAGAGAAATTAATAGAGAAACATCACCTCAGCTGC	1052	CTGCCCCCATGAAAGATGTGACAGTTTACAGATTTCCGTGATCCCGATGTTAGAAAATGGT	2132
1048	AGAAGAAATTTGAGCGGAGAGCTGAAGAGAAATTAATAGAGAAACATCACCTCAGCTGC	1107	CTGCCCCCATGAAAGATGTGACAGTTTACAGATTTCCGTGATCCCGATGTTAGAAAATGGT	2187
1053	ACCTTCCAAATGACCGAGTGGCGAATACTTGCCAGAGGTGAAACTTAAAAACAACAGGAT	1112	TCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCCTTGGGCATCCTTGGTGTGAATTTCTAGCA	2192
1108	ACCTTCCAAATGACCGAGTGGCGAATACTTGCCAGAGGTGAAACTTAAAAACAACAGGAT	1167	TCCAGAAAAACGAGCCCTCAGCTGGCGAATGCCCTTGGGCATCCTTGGTGTGAATTTCTAGCA	2247
1113	AGAGGAGCGGCTGAGGCGAGAGCTGCAAGAGCAATGCTGAGGACCCAGGAGAA	1172	AAATTCTACCAATATTTGCAATTTCTGAGCTAGCAAAATGTTCCAGTACATTTGGACCTAAACGG	2252
1168	AGAGGAGCGGCTGAGGCGAGAGCTGCAAGAGCAATGCTGAGGACCCAGGAGAA	1227	AAATTCTACCAATATTTGCAATTTCTGAGCTAGCAAAATGTTCCAGTACATTTGGACCTAAACGG	2307
1173	GAAGAGAGTCTGCAAGAAAGAAACATTTGAAAGAGATGAAGATGATGATGATGATGAT	1232	TGACTCTCATTTCTTTAAACAGATTTACAGTGAGCTGGCTTCACTCTCAGACCTTTATTTT	2312
1228	GAAGAGAGTCTGCAAGAAAGAAACATTTGAAAGAGATGAAGATGATGATGATGATGAT	1287	TGACTCTCATTTCTTTAAACAGATTTACAGTGAGCTGGCTTCACTCTCAGACCTTTATTTT	2367
1233	TGGGAAACATAGACCCCTACGTGGATGAAATCACTTAAAAACCAATGGCCATTTGAGATGG	1292	GCTTTGAGGTAATGCTGTTTGAACATTTTGGTGTGCACTGTGATCTCTGGGGAAGGCT	2427
1288	TGGGAAACATAGACCCCTACGTGGATGAAATCACTTAAAAACCAATGGCCATTTGAGATGG	1347	GCTTTGAGGTAATGCTGTTTGAACATTTTGGTGTGCACTGTGATCTCTGGGGAAGGCT	2472
1293	CCCATTTCTCAGGAGCGCACTGGACCAATGAAGATGATGAAGAGACTGCCCAA	1352	AGTCTTTTGTCTTCCAGCTAAGTGTGACCATTTTCTCTGGAACAATAAACATGT	2432
1348	CCCATTTCTCAGGAGCGCACTGGACCAATGAAGATGATGAAGAGACTGCCCAA	1407	AGTCTTTTGTCTTCCAGCTAAGTGTGACCATTTTCTCTGGAACAATAAACATGT	2487
1353	TCCTGAGGAATAAATCTTGATGAGCCAAATCGAGAAAGTGAATCAATATAGCAGCTC	1412	CTCTAAGCATTTGTTTCTTTGTGTGTGTGACATTAATAAGTCAATTTTGTGAATGAATAAT	2492
1408	TCCTGAGGAATAAATCTTGATGAGCCAAATCGAGAAAGTGAATCAATATAGCAGCTC	1467	CTCTAAGCATTTGTTTCTTTGTGTGTGTGACATTAATAAGTCAATTTTGTGAATGAATAAT	2547
1413	CTATGAACAATTTCAATGGTGAATTTGCCAATGGACACATTAATAATTTCCGAGTACAGTT	1472	ACTTTTCCCTTTGTGTTTGGCAGGTTTGTGTAATTTATTAAGAAATAATTTTGTAGCTGA	2552
1468	CTATGAACAATTTCAATGGTGAATTTGCCAATGGACACATTAATAATTTCCGAGTACAGTT	1527	ACTTTTCCCTTTGTGTTTGGCAGGTTTGTGTAATTTATTAAGAAATAATTTTGTAGCTGA	2607
1473	CCAGAGTTTTCACCTCGTTGTTTCTCTGGATCTTTAGAACCTGTGGCTGCGGCTCTGT	1532	GTACTATATAATTTACAACTTTAAGAAATTTATCAAGTTGGAACCAAGAAATA---GCNAG	2612
1528	CCAGAGTTTTCACCTCGTTGTTTCTCTGGATCTTTAGAACCTGTGGCTGCGGCTCTGT	1587	GTACTATATAATTTACAACTTTAAGAAATTTATCAAGTTGGAACCAAGAAATA---GCNAG	2664
1533	GCTTTCTGAGGATACACACTTACTGAGCAAGAGGAGAGCTCCATCCATGACAGAG	1592	GAAATGTACAAATTTTATCTCTGCGCAAGGAGACATCAATCTCTGATTTATAGTGTATGTAA	2672
1588	GCTTTCTGAGGATACACACTTACTGAGCAAGAGGAGAGCTCCATCCATGACAGAG	1647	GAAATGTACAAATTTTATCTCTGCGCAAGGAGACATCAATCTCTGATTTATAGTGTATGTAA	2724
1593	CAGAACGTTTTCAGCTTCCAGTACTGGGGATTTGCCAAAGCAAAAACCCGGGAGCTGA	1652	ATGCACCTGTAAATGTTTACTTTTGGATTAATAATATGGGA-GGGGAGCTCAAAATTCAGAAA	2783
1648	CAGAACGTTTTCAGCTTCCAGTACTGGGGATTTGCCAAAGCAAAAACCCGGGAGCTGA	1707	ATGCACCTGTAAATGTTTACTTTTGGATTAATAATATGGGA-GGGGAGCTCAAAATTCAGAAA	2783
1653	CTTGTGTGTGAATCCCTTGGATCCGCGGAATGCGAGATAAATTTAGAGTAAATTTGCTGA	1712	AGCTAAAAA 2741	
1708	CTTGTGTGTGAATCCCTTGGATCCGCGGAATGCGAGATAAATTTAGAGTAAATTTGCTGA	1767	AGCTACCAA 2792	
1713	CCTGGGAATGCTTGTGGTGCATAAACACTTCCAGCAAGACATCCAGACCGCTCAGTA	1772		
1768	CCTGGGAATGCTTGTGGTGCATAAACACTTCCAGCAAGACATCCAGACCGCTCAGTA	1827		

AC ABK51173;
XX 30-JUL-2002 (first entry)
DT cDNA encoding human cellular kinase SRPK-2 protein.
XX
DE Human; virucide; cytomegalovirus infection; CMV; cellular kinase; RICK;
KW RIP; Nck-Interacting kinase; MKK3; SRPK-2; gene; ss.
KW
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 188..2248
FT CDS
FT /*tag= a
FT /product= "Human cellular kinase SRPK-2"
XX
PN EP1201765-A2.
XX
XX 02-MAY-2002.
XX
XX 15-OCT-2001; 2001EP-00124604.
XX
XX 16-OCT-2000; 2000US-0240750P.
XX
XX (AXXI-) AXIMA PHARM AG.
XX
XX Schubart D, Habenberger P, Stein-Gerlach M, Bevec D;
PI
XX WPI; 2002-373930/41.
XX P-PSDB; AAU80373.
XX
XX Identifying agents for treatment or prevention of cytomegalovirus
PT infection, comprises contacting test compound with cellular kinase and
PT detecting change in cellular kinase activity.
XX
XX Disclosure; Page 39-43; 49pp; English.
XX
XX The present invention relates to a new method for identifying compounds
CC for treating and/or preventing cytomegalovirus (CMV) infection and/or
CC related diseases. The method of the invention comprises contacting a test
CC compound with at least one of the cellular kinases RICK, RIP, Nck-
CC Interacting kinase, MKK3 and SRPK-2 and detecting any change in kinase
CC activity. The method of the invention can be used to treat and/or prevent
CC CMV infections and related diseases. Oligonucleotides that can detect the
CC specified kinases can also be used for diagnosis of infection. The
CC present nucleic acid sequence encodes the human cellular kinase SRPK-2
CC protein of the invention, as described above
XX
XX Sequence 3745 BP; 1163 A; 719 C; 829 G; 1034 T; 0 U; 0 Other;
SQ
Query Match 76.8%; Score 2497; DB 6; Length 3745;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 20; Indels 10; Gaps 4;
QY 153 TCGGAAAAGCCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACCGCCACC 212
DB 217 TTCAGAAAGCCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACCGCCACC 276
QY 213 ACCACACACCCGACCTTTGCGAGACCCACACCCCGGAGCCAGAGGAGAGATCCT 272
DB 277 ACCACACACCCGACCTTTGCGAGACCCACACCCCGGAGCCAGAGGAGAGATCCT 336
QY 273 GGGATCAGATGATGAGGAGCAAGAGACCTCGGACTACTGCAAGGTGGATATCATCC 332
DB 337 GGGATCAGATGATGAGGAGCAAGAGACCTCGGACTACTGCAAGGTGGATATCATCC 396
QY 333 AGTGAATAATTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTTGGATGGG 392
DB 397 AGTGAATAATTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTTGGATGGG 456
QY 393 GCACCTTCTACTGTCTGCTGGGATATGACAGGGGAAAGATTTGTTGCAATGAA 452
DB 457 GCACCTTCTACTGTCTGCTGGGATATGACAGGGGAAAGATTTGTTGCAATGAA 516

QY 453 AGTTGTAAAAGTGCCTCCAGCATTTATACGAGACAGCCTTTGGATGAATAAATTTGCTCAA 512
DB 517 AGTTGTAAAAGTGCCTCCAGCATTTATACGAGACAGCCTTTGGATGAATAAATTTGCTCAA 576
QY 513 ATGTGTTTCGAGAAAGTGATCCCGAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA 572
DB 577 ATGTGTTTCGAGAAAGTGATCCCGAGTGACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA 636
QY 573 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCGATGGTCTTTCGAAAGTACTTGG 632
DB 637 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCGATGGTCTTTCGAAAGTACTTGG 696
QY 633 CCACCATCTCTCAAGTGGATCATCAAAATCCAACCTATCAAGGCCTCCAGTACGTTGTGT 692
DB 697 CCACCATCTCTCAAGTGGATCATCAAAATCCAACCTATCAAGGCCTCCAGTACGTTGTGT 756
QY 693 GAAAGAGTATCATTTTCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGCAGAT 752
DB 757 GAAAGAGTATCATTTTCGACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGCAGAT 816
QY 753 CATTCATCTGACATAAAGCCGGAATAATCTTTGATGTGTGGATGATGCATATGTGAG 812
DB 817 CATTCATCTGACATAAAGCCGGAATAATCTTTGATGTGTGGATGATGCATATGTGAG 876
QY 813 AAGAATGGCAGCTGAGGCCACCTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTTCAGGGTC 872
DB 877 AAGAATGGCAGCTGAGGCCACCTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTTCAGGGTC 933
QY 873 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTAAAAACAAAAAGAA 932
DB 934 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTAAAAACAAAAAGAA 993
QY 933 AAAAAGTGAAGAAAGAAAGAGAGGAGGAGGCTGAGTTATTTGGAGAGCGGCTGCGAGGAT 992
DB 994 AAAAAGTGAAGAAAGAAAGAGAGGAGGAGGCTGAGTTATTTGGAGAGCGGCTGCGAGGAT 1053
QY 993 AGAAGAAATTTGAGCGAGAGAGCTGAAAGGAAATATAGAGAAACATCATCCTCAGCTGC 1052
DB 1054 AGAAGAAATTTGAGCGAGAGAGCTGAAAGGAAATATAGAGAAACATCATCCTCAGCTGC 1113
QY 1053 ACCTTCCATATGACAGGAGTGGCGAATATCTGCCAGAGGTGAAACTTAAAAACAAACAGATT 1112
DB 1114 ACCTTCCATATGACAGGAGTGGCGAATATCTGCCAGAGGTGAAACTTAAAAACAAACAGATT 1173
QY 1113 AGAGAGGCGGCTGAGGAGAGAGCTGCAAGAGCAATGCTGAAGCTGAGGACGAGCAAGA 1172
DB 1174 AGAGAGGCGGCTGAGGAGAGAGCTGCAAGAGCAATGCTGAAGCTGAGGACGAGCAAGA 1233
QY 1173 GAAAGAAATGCTGAGAAAGAAACATTTGAAAAAGATGAAGATGATGTAGATCAGGAAT 1232
DB 1234 GAAAGAAATGCTGAGAAAGAAACATTTGAAAAAGATGAAGATGATGTAGATCAGGAAT 1293
QY 1233 TCGGAAATAGACCTTACGTGGATAGAAATCAGCTTAAACCAATGGCCATATTTAGAAATGG 1292
DB 1294 TCGGAAATAGACCTTACGTGGATAGAAATCAGCTTAAACCAATGGCCATATTTAGAAATGG 1353
QY 1293 CCATTTCTCAGCTGAGAGCAAACTGGAAGATGATGAAGAGAGCTGCCCCAAA 1352
DB 1354 CCATTTCTCAGCTGAGAGCAAACTGGAAGATGATGAAGAGAGCTGCCCCAAA 1413
QY 1353 TCCTGAGGAATATATCTTTGATGAGCCAAATCAGAAAGTGTATACATATAGCAGCTC 1412
DB 1414 TCCTGAGGAATATATCTTTGATGAGCCAAATCAGAAAGTGTATACATATAGCAGCTC 1473
QY 1413 CTATGAACAAATTCATGTGTGATTTGCCAAATGGACACATAAATTTCCGAGTCAAGTT 1472
DB 1474 CTATGAACAAATTCATGTGTGATTTGCCAAATGGACACATAAATTTCCGAGTCAAGTT 1533
QY 1473 CCAGAGTTTTCACCTCGTTGTTCTCTGGATCTCTTAGAACCTGTGGCTGCGGCTCTGT 1532
DB 1534 CCAGAGTTTTCACCTCGTTGTTCTCTGGATCTCTTAGAACCTGTGGCTGCGGCTCTGT 1593

CC hyperproliferating skin diseases, psoriasis, diabetes mellitus Type I
CC and/or II. Sequences of the invention are also useful as inhibitors for
CC mycobacterial protein kinases, nuclear export of TNF-alpha mRNA in TNF-
CC alpha mediated diseases, protein kinase and/or phosphatase and cellular
CC hyperproliferation of cancer cells. The invention is useful for treating
CC opportunistic infections (e.g. cancer, mycobacteria-induced opportunistic
CC infections), diabetes, inflammation, cancer cachexia, necrosis, gastric
CC ulcers, influenza, multiple sclerosis, neuropathic diseases, neuropathic
CC pain and polyneuropathy, neurodegenerative diseases (Alzheimer's disease,
CC Parkinson's disease), neurological disorder, gastrointestinal ulcerative,
CC Huntington's chorea, peripheral and/or central nerve diseases, atopic
CC eczema, psoriasis, urticaria, allergic reactions, rheumatoid arthritis,
CC osteoarthritis, ulcerative colitis, Crohn's disease, ischaemic diseases
CC and ischaemic heart disease, liver diseases and dysfunction of liver,
CC cardiovascular diseases, psychiatric disorders, schizophrenia, stroke,
CC alcoholism, attention deficit disorder, depression, obesity, learning
CC disabilities, pain, senile macular degeneration, diseases affecting the
CC immune system, impotence, male infertility, respiratory disorders, asthma
CC and infections. The present sequence is a human DNA related to the
CC invention
XX

SQ Sequence 3745 BP; 1163 A; 719 C; 829 G; 1034 T; 0 U; 0 Other;

Query Match 76.8%; Score 2497; DB 8; Length 3745;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 20; Indels 10; Gaps 4;
QY 153 TCGAAAAAGCCGAGCCTCAACAGAAAGCTCTTTAGTTCCTCTCCACCGCCACC 212
DB 217 TTCAGAAAGCCGAGCCTCAACAGAAAGCTCTTTAGTTCCTCTCCACCGCCACC 276
QY 213 ACCACACACCGCCACCTTGGCAGACCCACACCCCGGAGCCAGAGAGAGATCCT 272
DB 277 ACCACACACCGCCACCTTGGCAGACCCACACCCCGGAGCCAGAGAGAGATCCT 336
QY 273 GGGATCAGATGATGAGGACGACCTGGGACTACTGCAAGGTGGATATCATCC 332
DB 337 GGGATCAGATGATGAGGACGAGGACCTGGGACTACTGCAAGGTGGATATCATCC 396
QY 333 AGTGAANAATTGGAGACCTCTTCAATGGCCGGTATCATTTATAGAAAGCTTGGATGGG 392
DB 397 AGTGAANAATTGGAGACCTCTTCAATGGCCGGTATCATTTATAGAAAGCTTGGATGGG 456
QY 393 GCACCTCTCTACTGTCTGCTGTGCTGGATATGAGGGGAAAGATTTGTGCAATGAA 452
DB 457 GCACCTCTCTACTGTCTGCTGTGCTGGATATGAGGGGAAAGATTTGTGCAATGAA 516
QY 453 AGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCCTTGGATGAATAAATTTGCTCAA 512
DB 517 AGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCCTTGGATGAATAAATTTGCTCAA 576
QY 513 ATGTGTTTCGAGAAAGTGATCCCAAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATGA 572
DB 577 ATGTGTTTCGAGAAAGTGATCCCAAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATGA 636
QY 573 CGACTTCAAGATTTTCAAGCATGATGGGATACATGTCTGCATGGTCTTCAAGATCTGG 632
DB 637 CGACTTCAAGATTTTCAAGCATGATGGGATACATGTCTGCATGGTCTTCAAGATCTGG 696
QY 633 CCACCATCTCTCAAGTGATCATCAATCCAACTATCAAGCCTCCAGTACGTTGTGT 692
DB 697 CCACCATCTCTCAAGTGATCATCAATCCAACTATCAAGCCTCCAGTACGTTGTGT 756
QY 693 GAAGAGTATCATTCGACAGGCTCTTCAAGGGTTAGATTTACTTACAGATGAAGTGAAGAT 752
DB 757 GAAGAGTATCATTCGACAGGCTCTTCAAGGGTTAGATTTACTTACAGATGAAGTGAAGAT 816
QY 753 CATTTACTGACATTAAGCCGAAATATCTTTGATGTGTGGATGATGATATGTGAG 812
DB 817 CATTTACTGACATTAAGCCGAAATATCTTTGATGTGTGGATGATGATATGTGAG 876
QY 813 AAGATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCAGGTC 872

DB 877 AAGATGGCAGCTGAG---CCTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCAGGGTC 933
QY 873 TGCAGTGTAGTGGCTCCACAGCAGAAACCTATAGAAATAATCTATAAAACAAAAGAA 932
DB 934 TGCAGTGTAGTGGCTCCACAGCAGAAACCTATAGAAATAATCTATAAAACAAAAGAA 993
QY 933 AAAAAGTGAAGAAAGAAACAGAAAGCAGGCTGAGTTATTGGGAAGCGCTGAGGAGAT 992
DB 994 AAAAAGTGAAGAAAGAAACAGAAAGCAGGCTGAGTTATTGGGAAGCGCTGAGGAGAT 1053
QY 993 AGAAGAAATTTGAGCGAGAAAGCTGAAAGGAAATAATAGAAAGAAACATCACCTCAGCTGC 1052
DB 1054 AGAAGAAATTTGAGCGAGAAAGCTGAAAGGAAATAATAGAAAGAAACATCACCTCAGCTGC 1113
QY 1053 ACCTTCCAAATGACACGAGATGGGGAATACCTCCAGAGGTGAATCTAAAAACAACAGATT 1112
DB 1114 ACCTTCCAAATGACACGAGATGGGGAATACCTCCAGAGGTGAATCTAAAAACAACAGATT 1173
QY 1113 AGAGAGGCGGCTGAGGCGAGAGACTGCAAGAGCAATGTGAAGCTGAGGACCCAGGAAGA 1172
DB 1174 AGAGAGGCGGCTGAGGCGAGAGACTGCAAGAGCAATGTGAAGCTGAGGACCCAGGAAGA 1233
QY 1173 GAAAGAAATGCTGAGAAAGAAACAACTTGAAGAAAGATGAAGATGATGTAGATCAGGAAC 1232
DB 1234 GAAAGAAATGCTGAGAAAGAAACAACTTGAAGAAAGATGAAGATGATGTAGATCAGGAAC 1293
QY 1233 TCGGAACATAGACCTTACCTGATAGAAATACCTAAAAACAATGGCCATATTGAGAAATGG 1292
DB 1294 TCGGAACATAGACCTTACCTGATAGAAATACCTAAAAACAATGGCCATATTGAGAAATGG 1353
QY 1293 CCCATTCTCAGTGGAGCAGCAACTGGACGATGAAGATGATGAAGAGACTGCCCAA 1352
DB 1354 CCCATTCTCAGTGGAGCAGCAACTGGACGATGAAGATGATGAAGAGACTGCCCAA 1413
QY 1353 TCTGAGGAATAATCTTGTATGAGCCAAATGAGAAAGTGAATACACATATAGCAGCTC 1412
DB 1414 TCTGAGGAATAATCTTGTATGAGCCAAATGAGAAAGTGAATACACATATAGCAGCTC 1473
QY 1413 CTATGAACAAATTCATTTGTTGAATTCGCAAAATGGAAGCAATAAAATTCGAGTCAAGTT 1472
DB 1474 CTATGAACAAATTCATTTGTTGAATTCGCAAAATGGAAGCAATAAAATTCGAGTCAAGTT 1533
QY 1473 CCCAGAGTTTCCACCTGCTCTCTGATCTCTAGAACCTGCTGCGCTGCGGCTCTGT 1532
DB 1534 CCCAGAGTTTCCACCTGCTCTCTGATCTCTAGAACCTGCTGCGCTGCGGCTCTGT 1593
QY 1533 GCTTCTGAGGATCACCACTTACTGAGCAAGAGGAGAGCAGTCCATCCCATGACAGAAG 1592
DB 1594 GCTTCTGAGGATCACCACTTACTGAGCAAGAGGAGAGCAGTCCATCCCATGACAGAAG 1653
QY 1593 CAGAACGGTTTTCAGCTCCAGTACCTGGGATTTGCAAAAGCAAAACCCGGGCGAGCTGA 1652
DB 1654 CAGAACGGTTTTCAGCTCCAGTACCTGGGATTTGCAAAAGCAAAACCCGGGCGAGCTGA 1713
QY 1653 CTGTGTGGTGAATCCCTGGATCCCGGGAATGCAAGATAAATTTAGATGAATAAATTTGCTGA 1712
DB 1714 CTGTGTGGTGAATCCCTGGATCCCGGGAATGCAAGATAAATTTAGATGAATAAATTTGCTGA 1773
QY 1713 CTTGGGAATGCTTGTGTGGTGCATAAACACTTTCGGAAGACATCCAGACGGTCAAGTA 1772
DB 1774 CTTGGGAATGCTTGTGTGGTGCATAAACACTTTCGGAAGACATCCAGACGGTCAAGTA 1833
QY 1773 CCGCTCCATAGAGTTTAAATAGGAGCGGGGTACAGCACCCCTCGCGACATCTGGAGCAC 1832
DB 1834 CCGCTCCATAGAGTTTAAATAGGAGCGGGGTACAGCACCCCTCGCGACATCTGGAGCAC 1893
QY 1833 GCGGTGTATGGCATTTGAGCTGGCAACCGGAGATTTATTTGTTGAACCAATTTCTGGGA 1892
DB 1894 GCGGTGTATGGCATTTGAGCTGGCAACCGGAGATTTATTTGTTGAACCAATTTCTGGGA 1953
QY 1893 AGACTATTTCCAGAGAGAGAGACCATAGCCACATCATAGAGCTCTAGGCGAGTATTC 1952
DB 1954 AGACTATTTCCAGAGAGAGAGACCATAGCCACATCATAGAGCTCTAGGCGAGTATTC 2013

QY 1953 AAGGACATTTGCTCTATCTGGAATAATTTCTCGGAATTTCTCAATCGAGAGAGAACT 2012
DB 2014 AAGGACATTTGCTCTATCTGGAATAATTTCTCGGAATTTCTCAATCGAGAGAGAACT 2073
QY 2013 GCGACATACCAAGCTGAGCCCTGAGCCTCTTGATGATCTGTTGGAAGAATGATGG 2072
DB 2074 GCGACATACCAAGCTGAGCCCTGAGCCTCTTGATGATCTGTTGGAAGAATGATGG 2133
QY 2073 CTGGCCCATGATGATGCTGACAGTTTACAGATTTCTGATCCCGATGTTAGAAATGGT 2132
DB 2134 CTGGCCCATGATGATGCTGACAGTTTACAGATTTCTGATCCCGATGTTAGAAATGGT 2193
QY 2133 TCCAGAAAAACGAGCTCAGCTGCGGAATGCCCTCGGCATCTTGTTGAATTTCTTAGCA 2192
DB 2194 TCCAGAAAAACGAGCTCAGCTGCGGAATGCCCTCGGCATCTTGTTGAATTTCTTAGCA 2250
QY 2193 AATCTACCAATTTGCAATTTCTGAGCTAGCAATGTTCCAGTACATTTGGACCTAAACGG 2252
DB 2251 AATCTACCAATTTGCAATTTCTGAGCTAGCAATGTTCCAGTACATTTGGACCTAAACGG 2310
QY 2253 TGACTCTCATTTCTTAAACAGGATTTACAGTGAAGCTGGCTTCTATCTCAGACCTTTATTTT 2312
DB 2311 TGACTCTCATTTCTTAAACAGGATTTACAGTGAAGCTGGCTTCTATCTCAGACCTTTATTTT 2370
QY 2313 GCTTTGAGTACGTTGTTTGATGATTTGCTTTTGTGACGCTGATCTGCGGAAGGT 2372
DB 2371 GCTTTGAGTACGTTGTTGTTGATGATTTGCTTTTGTGACGCTGATCTGCGGAAGGT 2430
QY 2373 AGCTTTTGTCTCAGCTAGTATGTTTACTGACCATTTTCTTCTGGAACAATAACATGT 2432
DB 2431 AGCTTTTGTCTCAGCTAGTATGTTTACTGACCATTTTCTTCTGGAACAATAACATGT 2490
QY 2433 CTCTAAGCAATGTTCTGTTGTTGTGACATTTCAATGTCATTTTGTGATGAATAT 2492
DB 2491 CTCTAAGCAATGTTCTGTTGTTGTGACATTTCAATGTCATTTTGTGATGAATAT 2550
QY 2493 ACTTTCCCTCTGTTTGTGAGGTTTGTGACTATTTATGAGAAATATTTTACTGTA 2552
DB 2551 ACTTTCCCTCTGTTTGTGAGGTTTGTGACTATTTATGAGAAATATTTTACTGTA 2610
QY 2553 GTACTATATATTTCAATCTTAAGAAATTTATCAAGTTGGGAACCAAGAAATAGCAAG 2612
DB 2611 GTACTATATATTTCAATCTTAAGAAATTTATCAAGTTGGGAACCAAGAAATA---GCAAG 2667
QY 2613 GAAATGTACAAATTTTATCTTCTGCAAGGACATCATTTCTGATATATAGTATGTAA 2672
DB 2668 GAAATGTACAAATTTTATCTTCTGCAAGGACATCATTTCTGATATATAGTATGTAA 2727
QY 2673 ATGCACCTGTAAATGTTTACTTTGGATTAAATATGGAGGGGGGAGCTCAAAATTCAGAAA 2732
DB 2728 ATGCACCTGTAAATGTTTACTTTCCATTAATATGGGA---GGGGAGCTCAAAATTCAGAAA 2786
QY 2733 AGCTAAAAA 2741
DB 2787 AGCTACCAA 2795

RESULT 8
ID AAL54215 standard; DNA; 3745 BP.
XX AAL54215;
AC AAL54215;
XX 27-MAR-2003 (first entry)
DE SR protein-specific kinase-1 DNA, SEQ ID No 7.
KW Cytostatic; p53 pathway-modulating agent; SR protein; cancer; apoptosis;
KW lung cancer; breast cancer; serine/arginine protein-specific kinase;
KW colon cancer; ovarian cancer; SRPK; cell proliferation; human; ds.
XX Homo sapiens.
OS

XX WO200299427-A1.
XX 12-DEC-2002.
XX 03-JUN-2002; 2002WO-US017525.
XX 05-JUN-2001; 2001US-0296076P.
XX 10-OCT-2001; 2001US-0328605P.
XX 15-FEB-2002; 2002US-0357253P.
XX (EXEL-) EXELIXIS INC.
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-156865/15.
XX Identifying candidate p53 pathway-modulating agents useful as therapeutic
XX targets for disorders related with defective p53 function, by screening
XX for agents modulating serine/arginine protein-specific kinase activity.
XX Example; Page 111-113; 137pp; English.
XX The invention relates to a novel method for identifying candidate p53
XX pathway-modulating agents. The novel method comprises screening for
XX agents that modulate the activity of a serine/arginine protein-specific
XX kinase. The invention provides methods for utilizing p53 modifier genes
XX and polypeptides to identify candidate therapeutic agents that can be
XX used in the treatment of disorders associated with defective p53
XX function. The methods are also useful for modulating a p53 pathway in a
XX mammalian cell, or for diagnosing or treating a disease associated with
XX defective p53 function, e.g. cancers such as breast cancer, colon cancer,
XX lung cancer or ovarian cancer. The serine/arginine protein-specific
XX kinase (SRPK) polypeptides and nucleic acids are useful for identifying
XX and testing agents that modulate SRPK function. The animal models are
XX useful for in vivo assays to test the activity of a candidate p53-
XX modulating agent, or to assess the role of SRPK in a p53 pathway process
XX such as apoptosis or cell proliferation. This polynucleotide sequence
XX represents the DNA of a serine/arginine-rich domain protein-specific
XX kinase-1 of the invention
XX SQ Sequence 3745 BP; 1163 A; 719 C; 829 G; 1034 T; 0 U; 0 Other;
Query Match 76.8%; Score 2497; DB 8; Length 3745;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 20; Indels 10; Gaps 4;
QY 153 TCCGAAAAAGCCGAGAGCTCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACCGCCACC 212
DB 217 TTCAGAAAGCCGAGAGCTCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACCGCCACC 276
QY 213 ACCACCAACCGCCACCTTTTCCAGAGCCCCACACCCCGGAGCCAGAGGAGAGATCCT 272
DB 277 ACCACCAACCGCCACCTTTTCCAGAGCCCCACACCCCGGAGCCAGAGGAGAGATCCT 336
QY 273 GGGATCAGATGATGAGGAGCAAGAGAGCCCTGCGGACTACTGCGGACTACTGCGGACTACTGCGG 332
DB 337 GGGATCAGATGATGAGGAGCAAGAGAGCCCTGCGGACTACTGCGGACTACTGCGGACTACTGCGG 396
QY 333 AGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTGAAAGCTTGGATGGG 392
DB 397 AGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTGAAAGCTTGGATGGG 456
QY 393 GCATCTCTACTGCTGCTGCTGGGATATGAGGGGAAAGAAATTTGTTGCAATGAA 452
DB 457 GCATCTCTACTGCTGCTGCTGGGATATGAGGGGAAAGAAATTTGTTGCAATGAA 516
QY 453 AGTTGTAAAAAGTGGCCAGCATTTATACGAGACACCTTTGGATGAAATAAATTTCTCAA 512
DB 517 AGTTGTAAAAAGTGGCCAGCATTTATACGAGACACCTTTGGATGAAATAAATTTCTCAA 576
QY 513 ATGTGTTCCAGAAAGTATGATCCAGTCCCAACCAAGACATGTTGTTGCTCAGCTCATTTGA 572

Db	577	ATGTGTTGAGAAAGTGATCCAGTGAACCCAAACAAAGACATGGTGGTCCAGCTCATTTGA	636
Qy	573	CGACTTCAAGATTTTCAAGGCATGAATGGGATACATGCTGCGATGGTCTTTCGAAGTACTTGG	632
Db	637	CGACTTCAAGATTTTCAAGGCATGAATGGGATACATGCTGCGATGGTCTTTCGAAGTACTTGG	696
Qy	633	CCACCATCTCTCAAGTGGATCATCAAAATCCAACTATCAAGGCTCCCAAGTACTTGGTGTGT	692
Db	697	CCACCATCTCTCAAGTGGATCATCAAAATCCAACTATCAAGGCTCCCAAGTACTTGGTGTGT	756
Qy	693	GAAAGTATCATTTGACAGGTCCTTCAAGGTTTGAATTTACTTACACAGTAAGTGCAGAT	752
Db	757	GAAAGTATCATTTGACAGGTCCTTCAAGGTTTGAATTTACTTACACAGTAAGTGCAGAT	816
Qy	753	CATTTCATCTGACATAAAGCCGGAATAATCTTGTATGCTGTGTGGATGATGCATATGTGAG	812
Db	817	CATTTCATCTGACATAAAGCCGGAATAATCTTGTATGCTGTGTGGATGATGCATATGTGAG	876
Qy	813	AAGAATGGCAGCTTGAGGCCCATCTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCAAGGTC	872
Db	877	AAGAATGGCAGCTGAG--CCTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCAAGGTC	933
Qy	873	TGACGTGAGTACGGCTCCACAGCAGAAACCTATPAGGAAAAATATCTAAAAACAAAAAGAA	932
Db	934	TGACGTGAGTACGGCTCCACAGCAGAAACCTATPAGGAAAAATATCTAAAAACAAAAAGAA	993
Qy	933	AAAACTGAAAAAGAAACAGAGGCGAGCTGAGTTATTTGGAGAGCGCTCGCAGGAGAT	992
Db	994	AAAACTGAAAAAGAAACAGAGGCGAGCTGAGTTATTTGGAGAGCGCTCGCAGGAGAT	1053
Qy	993	AGAAGAAATTTGGAGCGAGAGCTGAAAGGAAAAATATAGAGAAAAACATCACTCTCAGCTGC	1052
Db	1054	AGAAGAAATTTGGAGCGAGAGCTGAAAGGAAAAATATAGAGAAAAACATCACTCTCAGCTGC	1113
Qy	1053	ACCTTCCAATGACCCAGGATGGCGAATPACTGCCAGAGGTGAAATCTAAAAACAAACAGGATT	1112
Db	1114	ACCTTCCAATGACCCAGGATGGCGAATPACTGCCAGAGGTGAAATCTAAAAACAAACAGGATT	1173
Qy	1113	AGAGAGGGCGGTGAGGCGAGAGCTGCAAAAGGCAATGTGTGAAGCTGAGGACAGAGAGA	1172
Db	1174	AGAGAGGGCGGTGAGGCGAGAGCTGCAAAAGGCAATGTGTGAAGCTGAGGACAGAGAGA	1233
Qy	1173	GAAAGAAAGTGTGAGAAAGAAAAATTCGAAAGAGATGAAGATGATGTAGATCAGGAATCT	1232
Db	1234	GAAAGAAAGTGTGAGAAAGAAAAATTCGAAAGAGATGAAGATGATGTAGATCAGGAATCT	1293
Qy	1233	TGCGAACATAGACCTTACGTGGATAGAATCACCTAAAAACCAATGGCCATATTGAGAAATGG	1292
Db	1294	TGCGAACATAGACCTTACGTGGATAGAATCACCTAAAAACCAATGGCCATATTGAGAAATGG	1353
Qy	1293	CCCATTTCTCATGGAGCAGCAACTGGACGATGAAGATGATGAAGAAAGTGCACCAAA	1352
Db	1354	CCCATTTCTCATGGAGCAGCAACTGGACGATGAAGATGATGAAGAAAGTGCACCAAA	1413
Qy	1353	TCCTGAGGAATATATCTTGTAGTACGCAATTCAGAAAGTGAATTAACATATAGCAGCTC	1412
Db	1414	TCCTGAGGAATATATCTTGTAGTACGCAATTCAGAAAGTGAATTAACATATAGCAGCTC	1473
Qy	1413	CTATGAACAATTAATGGTGAATTTGCCAAATGGACACATAAAATTTCCCGAGTTCACAGTT	1472
Db	1474	CTATGAACAATTAATGGTGAATTTGCCAAATGGACACATAAAATTTCCCGAGTTCACAGTT	1533
Qy	1473	CCGAGATTTTCCACTCGTGTGTCTCTGGATCTCTGAACTGTGGCTCTGGGGCTCTGT	1532
Db	1534	CCGAGATTTTCCACTCGTGTGTCTCTGGATCTCTGAACTGTGGCTCTGGGGCTCTGT	1593
Qy	1533	GCTTTCTGAGGATCACCACTTACTGAGCAAGGAGAGCAGTCCATCCCATGACAGAG	1592
Db	1594	GCTTTCTGAGGATCACCACTTACTGAGCAAGGAGAGCAGTCCATCCCATGACAGAG	1653
Qy	1593	CAGAACGGTTTTCAGCTCTCAGTACTGGGATTTGGCCAAAGCAAAACCCGGGAGCTGA	1652
Db	1654	CAGAACGGTTTTCAGCTCTCAGTACTGGGATTTGGCCAAAGCAAAACCCGGGAGCTGA	1713
Qy	1653	CTTGTGTGTAATCCCTGGATCCCGGAATGCGAGATAAAATTAGAGTAAAAATTTCTCTGA	1712
Db	1714	CTTGTGTGTAATCCCTGGATCCCGGAATGCGAGATAAAATTAGAGTAAAAATTTCTCTGA	1773
Qy	1713	CCTGGGAAATGCTTGTGTGGTGCATAAAACATCTTACCGAAGACATCCAGACGGCTCAGTA	1772
Db	1774	CCTGGGAAATGCTTGTGTGGTGCATAAAACATCTTACCGAAGACATCCAGACGGCTCAGTA	1833
Qy	1773	CCGCTCCATPAGAGTTTTAATAGGAGCGGGTACAGACCCCTCGGGAATCTTGGAGCAC	1832
Db	1834	CCGCTCCATPAGAGTTTTAATAGGAGCGGGTACAGACCCCTCGGGAATCTTGGAGCAC	1893
Qy	1833	GGCGGTATGSCATTTGAGCTGGCAACCGGAGATATTTGTTGTAACCAATTTCTGGGA	1892
Db	1894	GGCGGTATGSCATTTGAGCTGGCAACCGGAGATATTTGTTGTAACCAATTTCTGGGA	1953
Qy	1893	AGACTATTTCAGAGACGAAGACCAATAGCCCAACATCATAGAGCTGCTAGGAGTATTTCC	1952
Db	1954	AGACTATTTCAGAGACGAAGACCAATAGCCCAACATCATAGAGCTGCTAGGAGTATTTCC	2013
Qy	1953	AAGGCATTTGCTCTATCTGGAAAAATTTCTCGGAAATTTCTCAATCGCAGAGGAACT	2012
Db	2014	AAGGCATTTGCTCTATCTGGAAAAATTTCTCGGAAATTTCTCAATCGCAGAGGAACT	2073
Qy	2013	GGGACACATCACCAAGCTGAAGCCCTGGAGCCCTTTGATGATCTTGTGGAAGATATGG	2072
Db	2074	GGGACACATCACCAAGCTGAAGCCCTGGAGCCCTTTGATGATCTTGTGGAAGATATGG	2133
Qy	2073	CTGGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTCTGATCCCGATGTTAGAAATGGT	2132
Db	2134	CTGGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTCTGATCCCGATGTTAGAAATGGT	2193
Qy	2133	TCCAGAAAAACGAGCTCAGCTGGCGAAATGCTTGGCATCTCTGGTGAATTTCTTAGCA	2192
Db	2194	TCCAGAAAAACGAGCTCAGCTGGCGAAATGCTTGGCATCTCTGGTGAATTTCTTAGCA	2250
Qy	2193	AACTTCAACCAATTTGCAATTTCTGAGCTAGCAAAATGTTCCAGTACATTTGGAACCTAACCG	2252
Db	2251	AACTTCAACCAATTTGCAATTTCTGAGCTAGCAAAATGTTCCAGTACATTTGGAACCTAACCG	2310
Qy	2253	TGACTCTCATTTCTTAAACAGGATTTACAGTGAAGTGGCTTCACTCTCAGACCTTTATTTT	2312
Db	2311	TGACTCTCATTTCTTAAACAGGATTTACAGTGAAGTGGCTTCACTCTCAGACCTTTATTTT	2370
Qy	2313	GCTTTGAGTACTGTTGTTTGAATTTTGTGCACTGTGTGATCTCTGGGAGGCT	2372
Db	2371	GCTTTGAGTACTGTTGTTTGAATTTTGTGCACTGTGTGATCTCTGGGAGGCT	2430
Qy	2373	AGTCTTTTGTCTTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	2432
Db	2431	AGTCTTTTGTCTTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT	2490
Qy	2433	CTCTAAGCATTTCTTGT	2492
Db	2491	CTCTAAGCATTTCTTGT	2550
Qy	2493	ACTTTTCCCTTTGTTTGGCAGGTTTGTAACTATTATGAAGAAATATTTTACCTGA	2552
Db	2551	ACTTTTCCCTTTGTTTGGCAGGTTTGTAACTATTATGAAGAAATATTTTACCTGA	2610
Qy	2553	GTACTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2612
Db	2611	GTACTATATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	2667
Qy	2613	GAATGTACAAATTTTATCTTCTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2672
Db	2668	GAATGTACAAATTTTATCTTCTGGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2727
Qy	2673	ATGACCCCTGTAAATTTTACTTTTGAATTTAAATATGCGAGGGGGGAGCTCAAAATTTACAGAAA	2732
Db	2728	ATGACCCCTGTAAATTTTACTTTTGAATTTAAATATGCGAGGGGGGAGCTCAAAATTTACAGAAA	2786

[illegible]

Db 1354 ||||| CCCATTCCTCAGGACGCAACTGACCGATGAAGATGATGAAGAGACTGCCCCAA 1413
QY 1353 TCCTGAGGAATATAATCTTTGATGAGCCAAATGAGAAAGTGAATACACATATAGCAGCTC 1412
Db 1414 TCCTGAGGAATATAATCTTTGATGAGCCAAATGAGAAAGTGAATACACATATAGCAGCTC 1473
QY 1413 CTATGAACAATCAATGGTGAATTCGCAATGAGAGACATAAAATCCCGAGTCCAGATT 1472
Db 1474 CTATGAACAATCAATGGTGAATTCGCAATGAGAGACATAAAATCCCGAGTCCAGATT 1533
QY 1473 CCCAGAGTTTCCACCTCGTTCTCTGGATCCTTAGAACCTGTGGCCTGGGGCTCTGT 1532
Db 1534 CCCAGAGTTTCCACCTCGTTCTCTGGATCCTTAGAACCTGTGGCCTGGGGCTCTGT 1593
QY 1533 GCTTTCTGAGGATACCACTTACTGAGCAAGAGGAGAGCGTCCATCCCATGACAGAAG 1592
Db 1594 GCTTTCTGAGGATACCACTTACTGAGCAAGAGGAGAGCGTCCATCCCATGACAGAAG 1653
QY 1593 CAGAACGGTTTCAGCCTCCAGTACTGGGATTTGCCAAAGCAAAACCCGGGCGACTGA 1652
Db 1654 CAGAACGGTTTCAGCCTCCAGTACTGGGATTTGCCAAAGCAAAACCCGGGCGACTGA 1713
QY 1653 CTTGTGTGTGAATCCCTCGATCCGCGAATGCAGATATAAATTTAGAGTAAAAATTCCTGA 1712
Db 1714 CTTGTGTGTGAATCCCTCGATCCGCGAATGCAGATATAAATTTAGAGTAAAAATTCCTGA 1773
QY 1713 CTTGGGAATGCTTTGTGGTGATATAAACACTTCACGGAAGACATCCAGACCGCTCAGTA 1772
Db 1774 CTTGGGAATGCTTTGTGGTGATATAAACACTTCACGGAAGACATCCAGACCGCTCAGTA 1833
QY 1773 CCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGCACCCCTGGGACATCTGGAGCAC 1832
Db 1834 CCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGCACCCCTGGGACATCTGGAGCAC 1893
QY 1833 GCGGTGTATGGCAATTTAGCTGCGCAACGGGAGATTATTTGTTGAACCACTTCTGGGGA 1892
Db 1894 GCGGTGTATGGCAATTTAGCTGCGCAACGGGAGATTATTTGTTGAACCACTTCTGGGGA 1953
QY 1893 AGACTATTCCAGAGACGAAGACACATAGCCACATCATAGAGCTGCTAGGCAATATCC 1952
Db 1954 AGACTATTCCAGAGACGAAGACACATAGCCACATCATAGAGCTGCTAGGCAATATCC 2013
QY 1953 AAGGCATTTGCTCTATCTGGAATAATTTCTGGGAATTTCTCAATCGCAGAGGAATC 2012
Db 2014 AAGGCATTTGCTCTATCTGGAATAATTTCTGGGAATTTCTCAATCGCAGAGGAATC 2073
QY 2013 GCGACACATCACCAAGCTCAAGCCCTGGAGCCTCTTTGATGTACTTGTGAAAAGTATGG 2072
Db 2074 GCGACACATCACCAAGCTCAAGCCCTGGAGCCTCTTTGATGTACTTGTGAAAAGTATGG 2133
QY 2073 CTGGCCCCATGAAGATGCTGCAAGTTTACAGATTTCTCTGATCCCGATCTTAGAAATGGT 2132
Db 2134 CTGGCCCCATGAAGATGCTGCAAGTTTACAGATTTCTCTGATCCCGATCTTAGAAATGGT 2193
QY 2133 TCCAGAAAAACGAGCCTCAGCTGGGAAATGCTTGGGCATCTCTGGTGAATCTTAGCA 2192
Db 2194 TCCAGAAAAACGAGCCTCAGCTGGGAAATGCTTGGGCATCTCTGGTGAATCTTAGCA 2250
QY 2193 AATTCTACCAATATTGCAATCTGAGCTAGCAATGTTCCAGTACATTTGAGCTTAACCG 2252
Db 2251 AATTCTACCAATATTGCAATCTGAGCTAGCAATGTTCCAGTACATTTGAGCTTAACCG 2310
QY 2253 TGACTCTCAATCTTTAAACAGGATTAACAGTGAAGTGGCTTCTCCTCAGACCTTTATTTT 2312
Db 2311 TGACTCTCAATCTTTAAACAGGATTAACAGTGAAGTGGCTTCTCCTCAGACCTTTATTTT 2370
QY 2313 GCTTTGAGTACTGTTGTTGACATTTGCTTTTGTGCACTGTGATCTCGGGAGGGT 2372
Db 2371 GCTTTGAGTACTGTTGTTGACATTTGCTTTTGTGCACTGTGATCTCGGGAGGGT 2430
QY 2373 ACTCTTTCTCTCAGCTAAGTATTACTGACCATTTTCTTCTGGAACAATACATGT 2432
|||||

Db 2431 AGTCTTTTGTCTCTCAGCTAAGTAGTTTACTGACCATTTTCTTCTGGAACAATAACATGT 2490
QY 2433 CTCTAAGCAATTTGTTCTTGTGTGTGTGACATTTCAAAATGTCATTTTGTGAATGAAAAAT 2492
Db 2491 CTCTAAGCAATTTGTTCTTGTGTGTGTGACATTTCAAAATGTCATTTTGTGAATGAAAAAT 2550
QY 2493 AGTCTTCCCTTTGTGTTTGGCAGGTTTGTAACTATTTATGAAGAAATATTTTACCTGA 2552
Db 2551 ACTTTCCCTTTGTGTTTGGCAGGTTTGTAACTATTTATGAAGAAATATTTTACCTGA 2610
QY 2553 GTACTATATAATTTCAATCTTTAAGAAATTTATCAAGTTTGGGAACCAAGAAAAATAGCAAG 2612
Db 2611 GTACTATATAATTTCAATCTTTAAGAAATTTATCAAGTTTGGGAACCAAGAAAAATAGCAAG 2667
QY 2613 GAAATGTACAATTTTATCTTCTGGCAAGGACATCATTTCTCTGTATATAGTGTATGTAA 2672
Db 2668 GAAATGTACAATTTTATCTTCTGGCAAGGACATCATTTCTCTGTATATAGTGTATGTAA 2727
QY 2673 ATGCACCTCTGTAATTTTACTTTGGATTAAATATGAGGGGGGACTCAAAATTTTCAGAAA 2732
Db 2728 ATGCACCTCTGTAATTTTACTTTCCATTAAATATGGA--GGGGGACTCAAAATTTTCAGAAA 2786
QY 2733 AGCTAAAAA 2741
Db 2787 AGCTACCAA 2795
RESULT 10
AAL54217
ID AAL54217 standard; DNA; 2094 BP.
XX AC AAL54217;
XX DT 27-MAR-2003 (first entry)
XX SR protein-specific kinase-1 DNA, SEQ ID No 9.
DE Cytostatic; p53 pathway-modulating agent; SR protein; cancer; apoptosis;
KW lung cancer; breast cancer; serine/arginine protein-specific kinase;
KW colon cancer; ovarian cancer; SRPK; cell proliferation; human; ds.
XX Homo sapiens.
XX WO200299427-A1.
XX PD 12-DEC-2002.
XX 03-JUN-2002; 2002WO-US017525.
XX 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX (EXEL-) EXELIXIS INC.
XX Friedman L, Florman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-156865/15.
XX Identifying candidate p53 pathway-modulating agents useful as therapeutic
XX targets for disorders related with defective p53 function, by screening
XX for agents modulating serine/arginine protein-specific kinase activity.
XX Disclosure; Page 114-115; 137pp; English.
XX The invention relates to a novel method for identifying candidate p53
XX pathway-modulating agents. The novel method comprises screening for
XX agents that modulate the activity of a serine/arginine protein-specific
XX kinase. The invention provides methods for utilizing p53 modifier genes
XX and polypeptides to identify candidate therapeutic agents that can be
XX used in the treatment of disorders associated with defective p53
XX function. The methods are also useful for modulating a p53 pathway in a
XX mammalian cell, or for diagnosing or treating a disease associated with

CC defective p53 function, e.g. cancers such as breast cancer, colon cancer,
CC lung cancer or ovarian cancer. The serine/arginine protein-specific
CC kinase (SRPK) polypeptides and nucleic acids are useful for identifying
CC and testing agents that modulate SRPK function. The animal models are
CC useful for in vivo assays to test the activity of a candidate p53-
CC modulating agent, or to assess the role of SRPK in a p53 pathway process
CC such as apoptosis or cell proliferation. This polynucleotide sequence
CC represents the DNA of a serine/arginine-rich domain protein-specific
XX kinase-1 of the invention

SQ Sequence 2094 BP; 652 A; 458 C; 526 G; 458 T; 0 U; 0 Other;

Query Match 62.7%; Score 2039.6; DB 8; Length 2094;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2042; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	153	TCGGAAGAGCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACCGCCACC	212
DB	49	TTCAGAAAGCGCGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCCACCGCCACC	108
QY	213	ACCACACACCGCCACCTTTTCCAGACCCACACCCCGGAGCCAGAGGAGGATCCT	272
DB	109	ACCACACACCGCCACCTTTTCCAGACCCACACCCCGGAGCCAGAGGAGGATCCT	168
QY	273	GGGATCAGATGATGAGGACCAAGAGACCTTCGGACTACTGCNAAGGTGGATATCATCC	332
DB	169	GGGATCAGATGATGAGGACCAAGAGACCTTCGGACTACTGCNAAGGTGGATATCATCC	228
QY	333	AGTGAAGATTTGAGACCTCTCAATGGCCGGTATCATGTATTAGAAAGCTTGGATGGG	392
DB	229	AGTGAAGATTTGAGACCTCTCAATGGCCGGTATCATGTATTAGAAAGCTTGGATGGG	288
QY	393	GCACCTTCTACTGTCTGTCTGTCTGGATATGAGGGGAAAGATTTGTGCAATGAA	452
DB	289	GCACCTTCTACTGTCTGTCTGTCTGGATATGAGGGGAAAGATTTGTGCAATGAA	348
QY	453	AGTTGTAAAGTGGCCAGCATTTATACGGAGACAGCTTTGGATGAAATAAAATTCCTAA	512
DB	349	AGTTGTAAAGTGGCCAGCATTTATACGGAGACAGCTTTGGATGAAATAAAATTCCTAA	408
QY	513	ATGTGTTCCAGAAAGTGATCCAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATGGA	572
DB	409	ATGTGTTCCAGAAAGTGATCCAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATGGA	468
QY	573	CGACTTCAAGATTTCAAGGCATGAATGGGATACATGTCTCATGTGCTTCCGAAGTACTGG	632
DB	469	CGACTTCAAGATTTCAAGGCATGAATGGGATACATGTCTCATGTGCTTCCGAAGTACTGG	528
QY	633	CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCTCCCAAGTACGTTGTGT	692
DB	529	CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCTCCCAAGTACGTTGTGT	588
QY	693	GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTACTTACACAGTAAGTGAAGAT	752
DB	589	GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTACTTACACAGTAAGTGAAGAT	648
QY	753	CATTTCATCTGACATAAAGCCGAAATATCTTGTGTGTGGATGATGCATATGTGAG	812
DB	649	CATTTCATCTGACATAAAGCCGAAATATCTTGTGTGTGGATGATGCATATGTGAG	708
QY	813	AAGAAATGGCAGTGGAGCCCATGTAGTGGAGAAAGCAGGTGTCTCTCTCCCTTCAGGGTTC	872
DB	709	AAGAAATGGCAGTGGAGCCCATGTAGTGGAGAAAGCAGGTGTCTCTCTCTTCCTTCAGGGTTC	768
QY	873	TGCAGTGGAGTACGGCTCCACAGCAGAAACCTATAGGAAATAATCTAAACAAAGAA	932
DB	769	TGCAGTGGAGTACGGCTCCACAGCAGAAACCTATAGGAAATAATCTAAACAAAGAA	828
QY	933	AAACTGAAAGAAAGAAACAGAAAGGAGGCTGAGTTATTGGAGAGCGCTGCAGGAGAT	992
DB	829	AAACTGAAAGAAAGAAACAGAAAGGAGGCTGAGTTATTGGAGAGCGCTGCAGGAGAT	888
QY	993	AGAAGAAATTTGGAGCGGAGAGCTGAAAGGAAATAATAGAGAAACATCACCTCAGCTGC	1052

DB	889	AGAGAAATTTGGAGCGAGAGCTGAAAGGAAATAATAGAGAAACATCACCTCAGCTGC	948
QY	1053	ACCTTTCAATGACAGGATGGCGAATACTGCCAGAGGTGAAACTATAAAACAAACAGGATT	1112
DB	949	ACCTTTCAATGACAGGATGGCGAATACTGCCAGAGGTGAAACTATAAAACAAACAGGATT	1008
QY	1113	AGAGAGGCGGCTGAGCGAGAGCTGCAAAAGGCAATGGTGAAGCTGAGGACCAAGGAAGA	1172
DB	1009	AGAGAGGCGGCTGAGCGAGAGCTGCAAAAGGCAATGGTGAAGCTGAGGACCAAGGAAGA	1068
QY	1173	GAAAGAGATGCTCAGAAAGAAACATTTGAAAGATGAAAGATGAAAGATGAAAGATGAAAGAT	1232
DB	1069	GAAAGAGATGCTCAGAAAGAAACATTTGAAAGATGAAAGATGAAAGATGAAAGATGAAAGAT	1128
QY	1233	TGCGAAACATAGACCTCAGTGGATAGAAATCACCTTAAACCAATGGCCATATTGAGAATGG	1292
DB	1129	TGCGAAACATAGACCTCAGTGGATAGAAATCACCTTAAACCAATGGCCATATTGAGAATGG	1188
QY	1293	CCCATTTCTACTGAGCAGCAAACTGGACGATGAAAGATGATGAAGAGACTGCCCCAAA	1352
DB	1189	CCCATTTCTACTGAGCAGCAAACTGGACGATGAAAGATGATGAAGAGACTGCCCCAAA	1248
QY	1353	TCCTGAGGAATATTAATCTTGATGAGCCAAATGCGAAGATGATTACACATATAGCAGCTC	1412
DB	1249	TCCTGAGGAATATTAATCTTGATGAGCCAAATGCGAAGATGATTACACATATAGCAGCTC	1308
QY	1413	CTATGAACAAATTTCAATGGTGAATTTGCCAAATGGACGACATAAAATTTCCGAGTCA	1472
DB	1309	CTATGAACAAATTTCAATGGTGAATTTGCCAAATGGACGACATAAAATTTCCGAGTCA	1368
QY	1473	CCCGAGTTTTCCACCTGTTGTTCTCTGGAATCCTTAGAACCTGTGGCTCGGGCTCTGT	1532
DB	1369	CCCGAGTTTTCCACCTGTTGTTCTCTGGAATCCTTAGAACCTGTGGCTCGGGCTCTGT	1428
QY	1533	GCCTTCTGAGGATCCACCTTACTGACGAGAGAGAGAGAGCTCATCCCATGACAGAG	1592
DB	1429	GCCTTCTGAGGATCCACCTTACTGACGAGAGAGAGAGAGCTCATCCCATGACAGAG	1488
QY	1593	CAGAACGGTTTTCCAGCTCCTCAGTACTGGGGATTTTCCCAAAAGCAAAACCCGGGAGCTGA	1652
DB	1489	CAGAACGGTTTTCCAGCTCCTCAGTACTGGGGATTTTCCCAAAAGCAAAACCCGGGAGCTGA	1548
QY	1653	CTTGTGTGTAATCCCTGGAATCCCGGAATGCGAGATAAAATTTAGATGATAAAATTTGCTGA	1712
DB	1549	CTTGTGTGTAATCCCTGGAATCCCGGAATGCGAGATAAAATTTAGATGATAAAATTTGCTGA	1608
QY	1713	CCTGGAAATGCTTGTGGGTGCATAAACCTTCAACGGAAGACATCCAGACGCGTCAGTA	1772
DB	1609	CCTGGAAATGCTTGTGGGTGCATAAACCTTCAACGGAAGACATCCAGACGCGTCAGTA	1668
QY	1773	CCGCTTCCATAGAGGTTTTTAATAGGAGCGGGGTACAGCACCCCTCGCGACATCTGGAGCAC	1832
DB	1669	CCGCTTCCATAGAGGTTTTTAATAGGAGCGGGGTACAGCACCCCTCGCGACATCTGGAGCAC	1728
QY	1833	GGCGGTATGGCAATTTGAGCTGGCAACCGGAGATTTATTTGTTGACACCATTTCTGGGGA	1892
DB	1729	GGCGGTATGGCAATTTGAGCTGGCAACCGGAGATTTATTTGTTGACACCATTTCTGGGGA	1788
QY	1893	AGACTTTTCCAGAGAGCAAGACCAATAGCCACATCATAGAGCTGTCTAGGAGTATTC	1952
DB	1789	AGACTTTTCCAGAGAGCAAGACCAATAGCCACATCATAGAGCTGTCTAGGAGTATTC	1848
QY	1953	AAGGACATTTGCTTATCTTGGAAATAATTTCTGGAAATTTCTTCAATCCAGAGAGAACT	2012
DB	1849	AAGGACATTTGCTTATCTTGGAAATAATTTCTGGAAATTTCTTCAATCCAGAGAGAACT	1908
QY	2013	GGGACACATCACCAAGCTGAAGCCCTTGAGCTCTTTGATGTTCTGTTGGAAGTATGG	2072
DB	1909	GGGACACATCACCAAGCTGAAGCCCTTGAGCTCTTTGATGTTCTGTTGGAAGTATGG	1968
QY	2073	CTGGCCCCATGAAGATGCTGCACATTTACAGATTTCTCTGATCCCGATGTTTAGAAATGGT	2132

Db 1969 CTGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTCGATCCGATGTTAGATGCT 2028
QY 2133 TCCAGAAAAACGAGCCTCAGCTGGCGAATGCCTTCGGCATCCTTGGTTGAATCTTAGCA 2192
Db 2029 TCCAGAAAAACGAGCCTCAGCTGGCGAATGCCTTCGGCATCCTTGGTTGAATCTTAGCA 2088
QY 2193 AATTCT 2198
|||||
Db 2089 AATTCT 2094
|||||
RESULT 11
AAL54216
ID AAL54216 standard; DNA; 1956 BP.
XX
AC AAL54216;
XX
DT 27-MAR-2003 (first entry)
XX
DE SR protein-specific kinase-1 DNA, SEQ ID No 8.
XX
KW Cytostatic; p53 pathway-modulating agent; SR protein; cancer; apoptosis;
lung cancer; breast cancer; serine/arginine protein-specific kinase;
KW colon cancer; ovarian cancer; SRPK; cell proliferation; human; ds.
XX
OS Homo sapiens.
XX
PN WO200299427-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017525.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Friedman L, Plozman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
WPI; 2003-156865/15.
XX
XX
PT Identifying candidate p53 pathway-modulating agents useful as therapeutic
PT targets for disorders related with defective p53 function, by screening
PT for agents modulating serine/arginine protein-specific kinase activity.
XX
XX Disclosure; Page 113-114; 137pp; English.
XX
XX The invention relates to a novel method for identifying candidate p53
XX pathway-modulating agents. The novel method comprises screening for
XX agents that modulate the activity of a serine/arginine protein-specific
XX kinase. The invention provides methods for utilizing p53 modifier genes
XX and polypeptides to identify candidate therapeutic agents that can be
XX used in the treatment of disorders associated with defective p53
XX function. The methods are also useful for modulating a p53 pathway in a
XX mammalian cell, or for diagnosing or treating a disease associated with
XX defective p53 function, e.g. cancers such as breast cancer, colon cancer,
XX lung cancer or ovarian cancer. The serine/arginine protein-specific
XX kinase (SRPK) polypeptides and nucleic acids are useful for identifying
XX and testing agents that modulate SRPK function. The animal models are
XX useful for in vivo assays to test the activity of a candidate p53-
XX modulating agent, or to assess the role of SRPK in a p53 pathway process
XX such as apoptosis or cell proliferation. This polynucleotide sequence
XX represents the DNA of a serine/arginine-rich domain protein-specific
XX kinase-1 of the invention
XX
SQ Sequence 1956 BP; 612 A; 433 C; 486 G; 425 T; 0 U; 0 Other;
Query Match 53.9%; Score 1753.6; DB 8; Length 1956;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 1904; Conservative 0; Mismatches 4; Indels 138; Gaps 1;

QY 153 TCCGAAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCCCTCCACCGCCACC 212
|||
Db 49 TTCAGAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCCACCGCCACC 108
|||
QY 213 ACCACACACACCGCCACCTTTTCCAGACCCACACACCCCGGAGCCAGAGAGAGATCCT 272
|||
Db 109 ACCACACACCGCCACCTTTTCCAGACCCACACACCCCGGAGCCAGAGAGAGATCCT 168
|||
QY 273 GGGATCAGATGATGAGGAGCAAGAGACCTCTGGGACTACTGCAAAAGGTGGATATCATCC 332
|||
Db 169 GGGATCAGATGATGAGGAGCAAGAGACCTCTGGGACTACTGCAAAAGGTGGATATCATCC 228
|||
QY 333 AGTGAATAATTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAACCTTGGATGGG 392
|||
Db 229 AGTGAATAATTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAACCTTGGATGGG 288
|||
QY 393 GCACCTTCTACTGTCTGGCTGTGGATATGACAGGGGAAAAGATTTGTTGCAATGAA 452
|||
Db 289 GCACCTTCTACTGTCTGGCTGTGGATATGACAGGGGAAAAGATTTGTTGCAATGAA 348
|||
QY 453 AGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCCTTGGATGAAATAAATTTGCTCAA 512
|||
Db 349 AGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCCTTGGATGAAATAAATTTGCTCAA 408
|||
QY 513 ATGTGTTTCGAGAAAGTGTATCCAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATTTGA 572
|||
Db 409 ATGTGTTTCGAGAAAGTGTATCCAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATTTGA 468
|||
QY 573 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCATGGTCTTCAAGTACTTGG 632
|||
Db 469 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGTCATGGTCTTCAAGTACTTGG 528
|||
QY 633 CCACCATCTCTCAAGTGGATCATCAAATCAAATCAAGGCCTCCAGTAGTGTGTGTGT 692
|||
Db 529 CCACCATCTCTCAAGTGGATCATCAAATCAAATCAAGGCCTCCAGTAGTGTGTGTGT 588
|||
QY 693 GAAGAGTATCATTCGACAGGTCCTTCAAGGTTAGATTACTTACACAGTAAGTGCAGAT 752
|||
Db 589 GAAGAGTATCATTCGACAGGTCCTTCAAGGTTAGATTACTTACACAGTAAGTGCAGAT 635
|||
QY 753 CATTTCATCTGACATAAAGCCGGAATAATCTTGTGTGTGGATGATCATATGTGAG 812
|||
Db 636 ----- 635
|||
QY 813 AAGAATGGCAGCTGAGGCCATGAGTGGCAGAAAGCAGGTGCTCTCTCTCTTCAGGGTC 872
|||
Db 636 ----- 635
|||
QY 873 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAATAATCTTAAACCAAAAGAA 932
|||
Db 636 -----TGAGTACGGCTCCACAGCAGAAACCTATAGGAAATAATCTTAAACCAAAAGAA 690
|||
QY 933 AAAAAGTGAAGAAAGAAAGAGAGCAGGCTGAGTTATTGGAGAGCGGCTGCAGGAGAT 992
|||
Db 691 AAAAAGTGAAGAAAGAAAGAGAGCAGGCTGAGTTATTGGAGAGCGGCTGCAGGAGAT 750
|||
QY 993 AGAAGAATTTGGAGCAGAGAGCTGAAAGGAAATAATAGAAAGAAACATCACTTCAGCTGC 1052
|||
Db 751 AGAAGAATTTGGAGCAGAGAGCTGAAAGGAAATAATAGAAAGAAACATCACTTCAGCTGC 810
|||
QY 1053 ACCTTCCATGACCCAGGATGGCGAATATCTGCCAGAGGTGAAACTTAAACCAACAGGATT 1112
|||
Db 811 ACCTTCCATGACCCAGGATGGCGAATATCTGCCAGAGGTGAAACTTAAACCAACAGGATT 870
|||
QY 1113 ACAGAGAGCGGCTGAGGAGAGACATGCAAGAGCAATGCTGAAGCTGAGGACCCAGGAAGA 1172
|||
Db 871 AGAGAGAGCGGCTGAGGAGAGACATGCAAGAGCAATGCTGAAGCTGAGGACCCAGGAAGA 930
|||
QY 1173 GAAAGAAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGTAGATCAGGAAT 1232
|||
Db 931 GAAAGAAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGTAGATCAGGAAT 990
|||
QY 1233 TCCGAACATAGACCCCTACGTGGATAGATACACTTAAACCAATGGCCCATATTGAGAAATGG 1292
|||

483 GACAGCCTTGGATGAAATAAATTTGCTCAAATGTTTCGAGAAAGTGATCCAGTGACCC 542
| | | | |
1014 TGAAGCTGCGGACCTTGCAGTGAGTGTTCAGCTCTTAAAGGTGGCGCATCCAGAGTTGT 1073
| | | | |
543 AAACAAAGACATGGTGGTCCAGCTCATTTGACGATTTCAAGATTTTCAGGCAATGAATGGGAT 602
| | | | |
1074 TTGTTCCTCCGGTGGGTTTGGTGTCTGCTGACTTCAGGAATGAAGCCACAGACCTCTCA 1133
| | | | |
603 ACATGCTCGATGGTCTTTCGAAGTACTTTGGCCACCATCTCCTCAAGTGGATCATCAATC 662
| | | | |
1134 TGATGCTCGATGGTCTTTCGAAGTACTTTGGCCACCATCTCCTCAAGTGGATCATCAATC 1193
| | | | |
663 CAACTATCAAGGCTCCAGTACGTTGTGTAAGAGTATCATTTCCGACAGCTCTTCAAGS 722
| | | | |
1194 CAACTATCAAGGCTCCAGTACGTTGTGTAAGAGTATCATTTCCGACAGCTCTTCAAGS 1253
| | | | |
723 GTTAGATTACTTACACAGTAAGTCAAGATCAATTCATCTGACATATAAGCCGGAATAT 782
| | | | |
1254 GTTAGATTACTTACACAGTAAGTCAAGATCAATTCATCTGACATATAAGCCGGAATAT 1313
| | | | |
783 CTTGATGTTGGATGATGCATATGTGAGAGAAATGGCAGCTGAGGCCACTGAGTGGCA 842
| | | | |
1314 CTTGATGTTGGATGATGCATATGTGAGAGAAATGGCAGCTGAGGCCACTGAGTGGCA 1373
| | | | |
843 GAAAGCAGGTGCTCCTCCTCTTTCAGGCTCTGAGTGAAGTCCAGCAGCAGAAACC 902
| | | | |
1374 GAAAGCAGGTGCTCCTCCTCTTTCAGGCTCTGAGTGAAGTCCAGCAGCAGAAACC 1433
| | | | |
903 TATAGGAAATATCTTAAACAAACAAAGAAACCTGAAAGAAACAGAAAGAGGCGAGC 962
| | | | |
1434 TATAGGAAATATCTTAAACAAACAAAGAAACCTGAAAGAAACAGAAAGAGGCGAGC 1493
| | | | |
963 TGAGTTATTTGGAGAACGCTGAGGAGATAGAGAAATTTGGAGCGAGAAAGCTGAAAGGAA 1022
| | | | |
1494 TGAGTTATTTGGAGAACGCTGAGGAGATAGAGAAATTTGGAGCGAGAAAGCTGAAAGGAA 1553
| | | | |
1023 AATAATAGAGAAACATCACCTCAGCTCAGCTCCATCTCCATGATCCAGGATGGGAACTATG 1082
| | | | |
1554 AATAATAGAGAAACATCACCTCAGCTCAGCTCCATGATCCAGGATGGGAACTATG 1613
| | | | |
1083 CCCAGAGTGAAACTTAAACAAACAGGATTTAGAGGAGGCGGCTGAGGCGAGAGACTGCAAA 1142
| | | | |
1614 CCCAGAGTGAAACTTAAACAAACAGGATTTAGAGGAGGCGGCTGAGGCGAGAGACTGCAAA 1673
| | | | |
1143 GGACAAATGTGAGCTGAGGAGCCAGGAGAGAAAGATGCTGAGAGAAAGAAACATTTGA 1202
| | | | |
1674 GGACAAATGTGAGCTGAGGAGCCAGGAGAGAAAGATGCTGAGAGAAAGAAACATTTGA 1733
| | | | |
1203 AAAAGATGAAGATGATGATAGATCAGGAACCTTGGACATAGACCTTACGTTGGATGAGATC 1262
| | | | |
1734 AAAAGATGAAGATGATGATAGATCAGGAACCTTGGACATAGACCTTACGTTGGATGAGATC 1793
| | | | |
1263 ACCTAAACCAATGGCCATATTTAGAGAAATGGCCCATTTCTCACTGGAGCAGCAACTGGACGA 1322
| | | | |
1794 ACCTAAACCAATGGCCATATTTAGAGAAATGGCCCATTTCTCACTGGAGCAGCAACTGGACGA 1853
| | | | |
1323 TGAAGATGATGAGAGAGAGCTGCCCAAATCTTGAGGAATATAATCTTTGATGAGCCAAA 1382
| | | | |
1854 TGAAGATGATGAGAGAGAGCTGCCCAAATCTTGAGGAATATAATCTTTGATGAGCCAAA 1913
| | | | |
1383 TGCAGAAAGTGTATACATATAGCAGCTCCTATGACAAATTCATGGTGAATTTGCCAAA 1442
| | | | |
1914 TGCAGAAAGTGTATACATATAGCAGCTCCTATGACAAATTCATGGTGAATTTGCCAAA 1973
| | | | |
1443 TGGACGACATAAAATTTCCCGAGTTCACAGTTTCCAGAGTTTTCACACCTCGTTGTTCTCTGG 1502
| | | | |
1974 TGGACGACATAAAATTTCCCGAGTTCACAGTTTCCAGAGTTTTCACACCTCGTTGTTCTCTGG 2033
| | | | |
1503 ATCCTTTAGAACCTGTGGCTCTGTGCTTTCTGAGGGATCAACCTTACTGAGCA 1562
| | | | |
2034 ATCCTTTAGAACCTGTGGCTCTGTGCTTTCTGAGGGATCAACCTTACTGAGCA 2093
| | | | |
1563 AGAGGAGAGCAGTCCATCCATGACAGAGAGCAGACGGTTTCAGCCTCCAGTACTGGGA 1622
| | | | |

2094 AGAGGAGAGCAGTCCATCCATGACAGAAAGCAGACGGTTTCAGCCTCCAGTACTGGGA 2153
| | | | |
1623 TTTGCCAA----- 1630
| | | | |
2154 TTTGCCAAAGCGAGTAAATTTGATGAGCCGAGCCCTGCAACAGCCGACACTTCTGCAGC 2213
| | | | |
1631----- 1630
| | | | |
2214 CCCTTTGGTTCAGCAGTCTGGAATTTGCACGCCGACATGCCCTTCCATGGTGTCTGTAGA 2273
| | | | |
1631----- 1630
| | | | |
2274 GTGTCTTTCCAGTATGCTCACCCGATCAGTGTATTTCCCTGCTCACAGTTTCCACAGCTC 2333
| | | | |
1631----- 1630
| | | | |
2334 CCCATCATCTACAGCAGTGGTCTCCACAGTGGAGAGTGTACATCCCTGCTAATACCAGCAC 2393
| | | | |
1631----- 1630
| | | | |
2394 CATCAGGAAGCAAAACCCGGGCGAGCTGACTTTGTTGGTGAATCCCTCGATCCGCGAA 2453
| | | | |
1683 TGCAGATATAATTTAGAGTAAATAATGCTGACCTGGGAAATGCTTTGTTGGTGCATAAACA 1742
| | | | |
2454 TGCAGATATAATTTAGAGTAAATAATGCTGACCTGGGAAATGCTTTGTTGGTGCATAAACA 2513
| | | | |
1743 CTTCCGGAAGACATCCAGACGCTCAGTACCGCTCCATAGAGGTTTAAATAGGAGCGGG 1802
| | | | |
2514 CTTCCGGAAGACATCCAGACGCTCAGTACCGCTCCATAGAGGTTTAAATAGGAGCGGG 2573
| | | | |
1803 GTACAGCACCCCTGGCGCATCTGAGACACGCGGTGTATGGCAATTTGAGCTGGCAACGGG 1862
| | | | |
2574 GTACAGCACCCCTGGCGCATCTGAGACACGCGGTGTATGGCAATTTGAGCTGGCAACGGG 2633
| | | | |
1863 ACATTTATTTGTTTGAACACATCTTGGGGAAGACTATTTCCAGAGACGAAAGACCAATATG 1922
| | | | |
2634 ACATTTATTTGTTTGAACACATCTTGGGGAAGACTATTTCCAGAGACGAAAGACCAATATG 2693
| | | | |
1923 CCACATCATAGAGCTGCTAGGAGTATTTCCAGAGGCACTTTGCTCTATCTGGAATAATTTTC 1982
| | | | |
2694 CCACATCATAGAGCTGCTAGGAGTATTTCCAGAGGCACTTTGCTCTATCTGGAATAATTTTC 2753
| | | | |
1983 TCGGGAATTTCTCAATCCGAGAG 2005
| | | | |
2754 TCGGGAATTTCTCAATCCGAGAG 2776
| | | | |

RESULT 13

ADN98453

ID ADN98453 standard; cDNA; 1005 BP.

XX

AC ADN98453;

XX

XX 29-JUL-2004 (first entry)

XX

XX Novel human cDNA sequence #53.

DE

ds; gene; anti-inflammatory; dermatological; neuroprotective;
immunomodulator; antibacterial; virucide; antipsoriatic; cytostatic;
gene therapy; vaccine; inflammatory; CNS; immune disorder; cancer;
psoriasis; diabetes; early aging; hormonal imbalance;
ischemic heart disease; ulcerative colitis.

XX

OS Homo sapiens.

XX

XX WO2004038003-A2.

PN

XX 06-MAY-2004.

XX

XX 24-OCT-2003; 2003WO-US033947.

XX

XX 25-OCT-2002; 2002US-0421061P.

PR

PR 25-OCT-2002; 2002US-0421080P.
PR 25-OCT-2002; 2002US-0421552P.
PR 25-OCT-2002; 2002US-0421614P.
PR 30-OCT-2002; 2002US-0422177P.
PR 30-OCT-2002; 2002US-0422178P.
PR 15-NOV-2002; 2002US-0426155P.
PR 15-NOV-2002; 2002US-0426384P.
PR 15-NOV-2002; 2002US-0426394P.
PR 15-NOV-2002; 2002US-0426430P.
PR 15-NOV-2002; 2002US-0426916P.
PR 27-NOV-2002; 2002US-0429224P.
PR 27-NOV-2002; 2002US-0429275P.
PR 27-NOV-2002; 2002US-0429303P.
PR 27-NOV-2002; 2002US-0429328P.
PR 27-NOV-2002; 2002US-0429651P.
PR 04-DEC-2002; 2002US-0430645P.
PR 04-DEC-2002; 2002US-0430651P.
PR 04-DEC-2002; 2002US-0430657P.
PR 04-DEC-2002; 2002US-0430663P.
PR 04-DEC-2002; 2002US-0430668P.
PR 04-DEC-2002; 2002US-0430684P.
PR 05-DEC-2002; 2002US-0430937P.
PR 05-DEC-2002; 2002US-0430965P.
PR 05-DEC-2002; 2002US-0431458P.
PR 12-DEC-2002; 2002US-0433251P.
PR 12-DEC-2002; 2002US-0433500P.
PR 13-DEC-2002; 2002US-0433316P.
PR 13-DEC-2002; 2002US-0433318P.
PR 23-DEC-2002; 2002US-0436236P.
PR 03-JAN-2003; 2003US-0437914P.
PR 17-JAN-2003; 2003US-0440820P.
PR 17-JAN-2003; 2003US-0440821P.
PR 18-APR-2003; 2003US-0463700P.
PR 18-APR-2003; 2003US-0463708P.
PR 18-APR-2003; 2003US-0463716P.
PR 18-APR-2003; 2003US-0463732P.
PR 02-MAY-2003; 2003US-0467199P.
PR 02-MAY-2003; 2003US-0467201P.
PR 02-MAY-2003; 2003US-0467203P.
PR 02-MAY-2003; 2003US-0467230P.
PR 19-MAY-2003; 2003US-0471306P.
PR 19-MAY-2003; 2003US-0471336P.
PR 22-MAY-2003; 2003US-0472420P.
PR 22-MAY-2003; 2003US-0472430P.
PR 09-JUN-2003; 2003US-0476609P.
PR 09-JUN-2003; 2003US-0476621P.
PR 09-JUN-2003; 2003US-0476632P.
PR 09-JUN-2003; 2003US-0476641P.
PR 08-JUL-2003; 2003US-0485217P.
PR 08-JUL-2003; 2003US-0485218P.
PR 08-JUL-2003; 2003US-0485223P.
PR 08-JUL-2003; 2003US-0485224P.
PR 08-JUL-2003; 2003US-0485325P.
PR 08-JUL-2003; 2003US-0485359P.
PR 14-JUL-2003; 2003US-0486446P.
PR 14-JUL-2003; 2003US-0486480P.
PR 15-JUL-2003; 2003US-0486891P.
PR 15-JUL-2003; 2003US-0486960P.
PR 08-AUG-2003; 2003US-0493341P.
PR 08-AUG-2003; 2003US-0493370P.
PR 08-AUG-2003; 2003US-0493573P.
PR 08-AUG-2003; 2003US-0493577P.

(FIVE-) FIVE PRIME THERAPEUTICS INC.

PA Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
PI Halenbeck RF, Kothakota S, Lin H, Linnemann T, Pierce K, Wang Y;
PI Wong JGP, Wu G, Zhang H, Zeng C;
XX WPI; 2004-365511/34.
XX P-PSDB; ADN99237.

XX New nucleic acid molecules, useful in preparing a composition for

PT treating or preventing e.g. inflammatory, CNS, bacterial or viral
PT disorders, cancer, psoriasis, diabetes, ischemic heart disease or
XX ulcerative colitis.
XX Claim 1; SEQ ID NO 53; 532pp; English.
PS The invention relates to a nucleic acid molecule comprising a
CC polynucleotide sequence or its complement that encodes a polypeptide. The
CC nucleic acid is useful in preparing a composition for treating or
CC preventing inflammatory, CNS, immune, bacterial or viral disorder.
CC cancer, psoriasis, diabetes, early aging, hormonal imbalance, ischemic
CC heart disease or ulcerative colitis. This sequence corresponds to a
CC nucleic acid of the invention.

SQ Sequence 1005 BP; 312 A; 202 C; 254 G; 237 T; 0 U; 0 Other;

Query Match 21.9%; Score 712.8; DB 12; Length 1005;
Best Local Similarity 85.8%; Pred. No. 1.5e-150;
Matches 860; Conservative 0; Mismatches 87; Indels 55; Gaps 4;

QY 1119 GCGGCTGAGGCAGAGACTGCMAAGGACAATGGTGAAGCTGAGGACCAGGAAGAAGA 1178
DB 36 GGTAGCAGCAGATCCACGGCNAAGGACATGCTGAAGCTGAGGACCAGGAAGAAGA 95
QY 1179 AGATGCTGAGAAAGAAAACATTGAAAAGATGAAGATGATGTAGATCAGGAACCTGCGAA 1238
DB 96 AGACCTTGAGAAAGAAAACATTGAGAAA--GAAGATGATGTAGATCAGGAACCTGCGAA 152
QY 1239 CATAGACCTTACGTGGATAGATCACCTAAACCAATGGCCATATTCAGAAATGCCCATTT 1298
DB 153 CATAGACCTTACGTGGATAGATCACCTAAACCAATGGCCATCTTGAANAATGCCCATTT 212
QY 1299 CTCACTGGAGCAGCAACTGGACGATGAAGATGATGAAGAAGACTGCCCAAAATCCTGA 1358
DB 213 CTTACTGGAGCAGCAACTGGACGATGAAGAGGATAATGAAGAAGACCGCCCAATACTGA 272
QY 1359 GGAATATAATCTTGATGAGCCAAATGCAGAAAGTATACATATAGCAGCTCCTATGA 1418
DB 273 GGAATATAATCTTGATGAGCCAAATGCAGAAAGTATACATATAGCAGCTCCTATGA 332
QY 1419 ACAATTCATGCTGAATTTGCCAAATGGACGACATATAAATTTCCCGAGTCACAGTTCCCGA 1478
DB 333 ACAATTCATGCTGAATTTGCCAAATGGATGATGCAATATAAATTTCCCGAGTCACAGTTCCCGA 392
QY 1479 GTTTTCCACCTCGTTGTTCTTGATCCTTAGAACCTTGTTGGCCTCGGCTCTGTGCTTTC 1538
DB 393 GTTTTCCACCTCGTTGTTCTTGATCCTTAGAACCTTGTTGGCCTCGGCTCTGTGCTTTC 452
QY 1539 TGAGGGATCACCACTTACTGAGCAAGAGGAGGAGTCCATCCCATGACAGCAAGCAGAAC 1598
DB 453 TGAGGGATCACCACTTAC-----TGACAGAGCAGGAC 485
QY 1599 GGTTCAGCCTCCAGTACTGGGGATTTGCCAAAAGCAAAAACCCGGGAGCTGACTTGT 1658
DB 486 GGTTCAGCCTCCAGTACTGGGGATTTGCCAAAAGCAAAAACCCGGGAGGCTGACTTGT 545
QY 1659 GGTGAATCCCTGGATCCCGGAATGCAGATATAAATTAGTAGTAAATAATGCTGACCTGG 1718
DB 546 GATGAATCCCTGGATACACGGAATGCAGATAAATAATTAGTAGTAAATAATGCTGACCTGG 605
QY 1719 AAATGCTTGTGGTGATCAACCTTCCAGGAAGACATCCAGAGCGCTCAGTACCGCTC 1778
DB 606 AAATGCTTGTGGTGATCAACCTTCCAGGAAGACATCCAGATGATCAATACCGTTC 665
QY 1779 CATAGAGGTTTTAATAGGAGGGGTACAGCAACCTTCGGGACATCTGGAGACCGGCTG 1838
DB 666 CATAGAGGTTTTAATAGGAGGGGTACAGCAACCTTCGGGACATCTGGAGACCGGCTG 725
QY 1839 ATAGGCAATTTAGCTGGCAACCGGAGATTTTGTGTAACACCATTTCTGGGGAAGACTA 1898
DB 726 ATAGGCAATTTAGCTGGCAATCGGCAATTTTGTGTAACACCATTTCTGGGGAAGACTA 785
QY 1899 TTCCAGAGACGAAGACCACATAGCCCATCATAGCTGCTAGGACG----- 1946

Qy 1549 CCACCTTACTGAGCAAGAGGAGCAGTCCATCCCATGACAGAGCAGACGGTTTCAGCC 1608
 Db |||||
 475 CCACCTTAC-----TGCAGAGCAGGACGGTTTCAGCC 507
 Qy 1609 TCCAGTACTGGGGATTGTGCAAAAGCAAAACCCGGGCAGCTGACTTGTGTGGTGAATCCC 1668
 Db |||||
 508 TCCAGTACTGGGGATTGTGCAAAAGCAAAACCCGGGCAGCTGACTTGTGTGAATCCC 567
 Qy 1669 CTGGATCCCGGGAATGCAGATAAAATTAGAGTAAAAATGCTGACCTGGGAAATGCTTGT 1728
 Db |||||
 568 CTGGATACACGGAATGCAGATAAAATTAGAGTAAAAATGCTGACCTGGGAAATGCTTGT 627
 Qy 1729 TGGGTGCATATAACACTTCCAGGAGACATCCAGACCGCTCAGTACCGCTCCATAGAGGTT 1788
 Db |||||
 628 TGGGCCCATAAACACTTCCAGGAGACATCCAGATGAGTCAATACCGTTCCATAGAGGTT 687
 Qy 1789 TTAATAGGAGCGGGGTACAGCACCCCTGCGGCACATCTGGAGCACGGCGTGTATGGCATTT 1848
 Db |||||
 688 TTAATAGGAGAGGGATACAGCACCCCTGCGGCACATCTGGAGCACGGCGTGTATGGCATTT 747
 Qy 1849 GAGCTGGCAACGGGAGATTATTGTTTGAACCAACATTTCTGGGGAAGACTATTCCAGAGAC 1908
 Db |||||
 748 GAGCTGGCAATGGCGGATTATTGTTTGAACCAACATTTCTGGGGAAGACTATTCCAGAGAC 807
 Qy 1909 GAAGACCACATAGCCCAACATCATAGAGCTGCTAGGCAG----- 1946
 Db |||||
 808 GAAGATCACATAGCATTTGATCATTTGAACCTGCTGGGGAGAGTCCGGTTTACTGACCGTTTT 867
 Qy 1947 -TATTCGAAGGCATTTGCTCTATCTGGAATAATTTCTCGGAAATTTCTCAATCGCAGAG 2005
 Db |||||
 868 CTTCTGGAATAAATAACGCTATGTTGGGAAATATTTTCAGAGAGTTTTTCCAGAAAAG 927
 Qy 2006 GAGAACTGCGACACATCACCAAGCTGAAGCCCTGGAGCCCTTTTGATGTACTTTGTGGA 2065
 Db |||||
 928 GAGAACTGCG--ACATCACCAAGCTGAACCCCTGGTGCCCTTTTGATGTACTTTGTGGA 985
 Qy 2066 AGTATGGCTGGCCCCCATGAAGATGCTGCACAG 2097
 Db |||||
 986 AGTATGGCTGGCCCCCATGAAGATGCTGCATAG 1017

Search completed: January 23, 2005, 06:36:48
 Job time : 1478.41 secs

This Page Blank (uspio)

BEST AVAILABLE COPY

241 CCCACACCCCGAGCCAGAGGAGGAGATCCTGGGATCAGATGATGAGGAGCAAGAGAC 300
Db
241 CCCACACCCCGAGCCAGAGGAGGAGATCCTGGGATCAGATGATGAGGAGCAAGAGAC 300
Qy
301 CCTGGGACTACTGCAAAAGGTGGATATCATCCAGTGAAATTTGGAGACTCTTCAATGGC 360
Db
301 CCTGGGACTACTGCAAAAGGTGGATATCATCCAGTGAAATTTGGAGACTCTTCAATGGC 360
Qy
361 CGGTATCATGTTATTAGAAAGCTTCGATGGGGGCACTTCTCTACTGCTCTGGCTGGCTGG 420
Db
361 CGGTATCATGTTATTAGAAAGCTTCGATGGGGGCACTTCTCTACTGCTCTGGCTGGCTGG 420
Qy
421 GATATCGCGGGAAGAATTTCTGCAATGAAAGTTGTAAAGTGCCTCCAGCATTTATACG 480
Db
421 GATATCGCGGGAAGAATTTCTGCAATGAAAGTTGTAAAGTGCCTCCAGCATTTATACG 480
Qy
481 GAGACAGCCTTGGATGAAATAAAATTTGCTCAAAATGTTTCGAGAAAGTGATCCCAAGTAC 540
Db
481 GAGACAGCCTTGGATGAAATAAAATTTGCTCAAAATGTTTCGAGAAAGTGATCCCAAGTAC 540
Qy
541 CCAAAACAAGACATGCTGCTCCAGCTCATTTGACGACTTCAAGATTTTCAGGCAATGAATGG 600
Db
541 CCAAAACAAGACATGCTGCTCCAGCTCATTTGACGACTTCAAGATTTTCAGGCAATGAATGG 600
Qy
601 ATACATGCTGCATGCTTCCGAAGTACTTGGCCACCCTCTCCTCAAGTGGATCAATAA 660
Db
601 ATACATGCTGCATGCTTCCGAAGTACTTGGCCACCCTCTCCTCAAGTGGATCAATAA 660
Qy
661 TCAACTATCAAGGCTCCAGTACGTTGTGTGAAGATATCATTTGCAAGTCTTCA 720
Db
661 TCCAACTATCAAGGCTCCAGTACGTTGTGTGAAGATATCATTTGCAAGTCTTCA 720
Qy
721 GGGTTAGATTACTTACACAGTAAAGTGAAGATCATTTACATGACATTAAGCCGGAAT 780
Db
721 GGGTTAGATTACTTACACAGTAAAGTGAAGATCATTTACATGACATTAAGCCGGAAT 780
Qy
781 ATCTTGATGTTGGATGATGATATGTAAGAAATGGCAGTGAAGCCACTGAGTGG 840
Db
781 ATCTTGATGTTGGATGATGATATGTAAGAAATGGCAGTGAAGCCACTGAGTGG 840
Qy
841 CAGAAAGCAGGCTCCTCCTCTTCCAGGCTTCGAGTGAAGTACGCTCCACAGCAGAA 900
Db
841 CAGAAAGCAGGCTCCTCCTCTTCCAGGCTTCGAGTGAAGTACGCTCCACAGCAGAA 900
Qy
901 CTTATAGGAAAATATCTAAAAACAAAAAGAAAACTGAAAAAGAAACAGAAAGGCGAG 960
Db
901 CTTATAGGAAAATATCTAAAAACAAAAAGAAAACTGAAAAAGAAACAGAAAGGCGAG 960
Qy
961 GCTGAGTTATTGGAGAGCGCTGCAGAGATAGAAAGAAATTTGGAGGAGAACTGAAAGG 1020
Db
961 GCTGAGTTATTGGAGAGCGCTGCAGAGATAGAAAGAAATTTGGAGGAGAACTGAAAGG 1020
Qy
1021 AATAATAGAAAGAAACATCCTCAGCTGCACCTTCCAAATGACCCAGGATGCGGAATAC 1080
Db
1021 AATAATAGAAAGAAACATCCTCAGCTGCACCTTCCAAATGACCCAGGATGCGGAATAC 1080
Qy
1081 TGCCACAGAGTGAAACTAAAAACAACAGGATTTAGAGAGCGGCTGAGGACAGACTGCA 1140
Db
1081 TGCCACAGAGTGAAACTAAAAACAACAGGATTTAGAGAGCGGCTGAGGACAGACTGCA 1140
Qy
1141 AAGGACATGTTGAGAGCTGAGGACAGGAGAGAAAGATGCTGAGAAAGAAACATT 1200
Db
1141 AAGGACATGTTGAGAGCTGAGGACAGGAGAGAAAGATGCTGAGAAAGAAACATT 1200
Qy
1201 GAAAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db
1201 GAAAGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Qy
1261 TCACTTAAACCAATGGCCATTTGAGAAATGGCCATTTCTCACTGAGCAGCAACTGGAC 1320
Db
1261 TCACTTAAACCAATGGCCATTTGAGAAATGGCCATTTCTCACTGAGCAGCAACTGGAC 1320
Qy
1321 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380

1321 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Qy
1381 AATCAGAGAGTGAATACATATAGCAGCTCTCTATGAAACAAATTTCAATGGTGAATGGCA 1440
Db
1381 AATCAGAGAGTGAATACATATAGCAGCTCTCTATGAAACAAATTTCAATGGTGAATGGCA 1440
Qy
1441 AATGAGACACATAAAATTTCCGAGTCAAGTTCACAGATTTTCCAGCTCTGCTTCTCT 1500
Db
1441 AATGAGACACATAAAATTTCCGAGTCAAGTTCACAGATTTTCCAGCTCTGCTTCTCT 1500
Qy
1501 GGATCTTTAGAACTGCTGGCTGGCTCTGCTGCTTCTGAGGATCAACACTTACTGAG 1560
Db
1501 GGATCTTTAGAACTGCTGGCTGGCTCTGCTGCTTCTGAGGATCAACACTTACTGAG 1560
Qy
1561 CAAGAGAGAGCAGTCCATCCATGACAGAGAGAGAGCGGTTTCCAGCTCTCAGTACTGGG 1620
Db
1561 CAAGAGAGAGCAGTCCATCCATGACAGAGAGAGAGCGGTTTCCAGCTCTCAGTACTGGG 1620
Qy
1621 GATTTGCCAAAGCAAAACCCCGGCGAGCTGACTTGTGTGGAATTCCTCTGGATCCGCGG 1680
Db
1621 GATTTGCCAAAGCAAAACCCCGGCGAGCTGACTTGTGTGGAATTCCTCTGGATCCGCGG 1680
Qy
1681 AATCAGATATAAATTTAGATGATAAATTTGCTGACCTGGGAAATGCTTGTGGGTGATATA 1740
Db
1681 AATCAGATATAAATTTAGATGATAAATTTGCTGACCTGGGAAATGCTTGTGGGTGATATA 1740
Qy
1741 CACTTCCGAGAGACATCCAGACGCTCAGTACCGCTCCATAGAGGTTTTAATAGGAGCG 1800
Db
1741 CACTTCCGAGAGACATCCAGACGCTCAGTACCGCTCCATAGAGGTTTTAATAGGAGCG 1800
Qy
1801 GGGTACAGACCCCTGCGGACATCTGGAGCAGCGGTGTATGGCATTTGAGCTGGCAACG 1860
Db
1801 GGGTACAGACCCCTGCGGACATCTGGAGCAGCGGTGTATGGCATTTGAGCTGGCAACG 1860
Qy
1861 GGAGATTAATTTGTTGAAACCACTTCTGGGGAAGACTATTTCCAGAGAGAGAGACACATA 1920
Db
1861 GGAGATTAATTTGTTGAAACCACTTCTGGGGAAGACTATTTCCAGAGAGAGAGACACATA 1920
Qy
1921 GCCCACATCATAGAGCTGCTAGGCAATTTCCAAAGGCACTTTGCTCTATCTGGAATAAT 1980
Db
1921 GCCCACATCATAGAGCTGCTAGGCAATTTCCAAAGGCACTTTGCTCTATCTGGAATAAT 1980
Qy
1981 TCTCGGAAATTTCTCAATCGCAGAGAGAACTGGGACACATCAACAGCTCAAGCCCTGG 2040
Db
1981 TCTCGGAAATTTCTCAATCGCAGAGAGAACTGGGACACATCAACAGCTCAAGCCCTGG 2040
Qy
2041 AGCCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db
2041 AGCCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Qy
2101 ACAGATTTCTGATCCCGATGTAAGAAATGGTTCAGAAAAACGAGCTCAGCTGGCGAA 2160
Db
2101 ACAGATTTCTGATCCCGATGTAAGAAATGGTTCAGAAAAACGAGCTCAGCTGGCGAA 2160
Qy
2161 TGCTTCCGATCCTTGGTGAATTTCTAGCAATTTCTACCAATTTCTAGCTTCTGAGCTA 2220
Db
2161 TGCTTCCGATCCTTGGTGAATTTCTAGCAATTTCTACCAATTTCTAGCTTCTGAGCTA 2220
Qy
2221 GCAAAATGTTCCAGTACATTTGACCTAAACCGTGAATCTCTCTTTTAAACAGGATTA 2280
Db
2221 GCAAAATGTTCCAGTACATTTGACCTAAACCGTGAATCTCTCTTTTAAACAGGATTA 2280
Qy
2281 GTGAGCTGGCTTCACTCTCAGACCTTTATTTGCTTTGAGGTAATGTTGTTGATTTT 2340
Db
2281 GTGAGCTGGCTTCACTCTCAGACCTTTATTTGCTTTGAGGTAATGTTGTTGATTTT 2340
Qy
2341 GCTTTTGTGACTGTGATCCTGGGAGAGGAGTCTTTTGTCTTCACTTCACTTCACTTCA 2400
Db
2341 GCTTTTGTGACTGTGATCCTGGGAGAGGAGTCTTTTGTCTTCACTTCACTTCACTTCA 2400
Qy
2401 CTGACATTTTCTTCTGGAACCAATAACATGCTCTTAAAGCATTTCTTGTGTTGTG 2460

Db 2401 CTGACCATTTCTCTGGAAACAATAACATGTCTCTAAGCATGTTTCTTGTTGTTGTG 2460
QY 2461 ACATTCAAATGTCATTTTGTGAATGAATAAATCTTCCCTTTGTTGTTTGGCAGGTTT 2520
Db 2461 ACATTCAAATGTCATTTTGTGAATGAATAAATCTTCCCTTTGTTGTTTGGCAGGTTT 2520
QY 2521 TGTAACTATTATGAAGAATAATTTTGTAGCTGTAGTACTATATATATTTACAATCTTAAGAAA 2580
Db 2521 TGTAACTATTATGAAGAATAATTTTGTAGCTGTAGTACTATATATATTTACAATCTTAAGAAA 2580
QY 2581 TTATCAAGTTGGGAACCAAGAAATAGCAAGGAAATGTACAAATTTTATCTTCTGGCAAA 2640
Db 2581 TTATCAAGTTGGGAACCAAGAAATAGCAAGGAAATGTACAAATTTTATCTTCTGGCAAA 2640
QY 2641 GGGACATCATCTCTGTATTATAGTATGTAATGACACCTGTAAATGTACTTTGGATT 2700
Db 2641 GGGACATCATCTCTGTATTATAGTATGTAATGACACCTGTAAATGTACTTTGGATT 2700
QY 2701 AAATATGGGAGGGGAGCTCAAAATTTTCAGAAAGCTTAAAGGAAAGGAAAGGAAAGG 2760
Db 2701 AAATATGGGAGGGGAGCTCAAAATTTTCAGAAAGCTTAAAGGAAAGGAAAGGAAAGG 2760
QY 2761 AAAAATACCTTATATATAAATAACCTTTCTTTGTTTGTGTTTCTTATTTTCATATT 2820
Db 2761 AAAAATACCTTATATATAAATAACCTTTCTTTGTTTGTGTTTCTTATTTTCATATT 2820
QY 2821 ATTAATACACTTAACTGTCGGAAGACATATGAAAGGAAATTAATACATGAAAGGATCA 2880
Db 2821 ATTAATACACTTAACTGTCGGAAGACATATGAAAGGAAATTAATACATGAAAGGATCA 2880
QY 2881 AAAATCATAAATCAAAACCCCACTATAGTCCACGACAAATTCATCTCGGCGGTCACTT 2940
Db 2881 AAAATCATAAATCAAAACCCCACTATAGTCCACGACAAATTCATCTCGGCGGTCACTT 2940
QY 2941 TTTAACTATTATCTAGTACCTGAGACTCTGTGTCTCAATATTAATATTTCTAAATCTAC 3000
Db 2941 TTTAACTATTATCTAGTACCTGAGACTCTGTGTCTCAATATTAATATTTCTAAATCTAC 3000
QY 3001 CACCAAGTTAGGCCCGGTAAATGTCGTCTCTCGTGAATCTGTCATACATACATATTTTCT 3060
Db 3001 CACCAAGTTAGGCCCGGTAAATGTCGTCTCTCGTGAATCTGTCATACATACATATTTTCT 3060
QY 3061 ATTATTTAGTGGTCTGTTTATCTTTTCCGCGACATCTTTGTTCACTATTTTCTAGTTA 3120
Db 3061 ATTATTTAGTGGTCTGTTTATCTTTTCCGCGACATCTTTGTTCACTATTTTCTAGTTA 3120
QY 3121 CTCTATCTTTGGGCTGATTAACTCTCTCATTTATCTCATATAAACTTCTGAATTTTTC 3180
Db 3121 CTCTATCTTTGGGCTGATTAACTCTCTCATTTATCTCATATAAACTTCTGAATTTTTC 3180
QY 3181 ACATAAACTACTAGAGTACCTCACCATCTCTGTTTTTAAACGCGAGGATTTACTATATA 3240
Db 3181 ACATAAACTACTAGAGTACCTCACCATCTCTGTTTTTAAACGCGAGGATTTACTATATA 3240
QY 3241 ATTACTATTATAA 3253
Db 3241 ATTACTATTATAA 3253

RESULT 2
AX666205
LOCUS AX666205 3253 bp DNA linear PAT 26-MAR-2003
DEFINITION Sequence 1 from Patent WO02057458.
ACCESSION AX666205
VERSION AX666205.1 GI:29291001
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Yan, C., Abu-Threideh, J., Shao, W., Merklöv, G., di Francesco, V. and Beasley, E.M.

TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: WO 02057458-A 1 25-JUL-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
Source 1. 3253
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 100.0%; Score 3253; DB 6; Length 3253;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3253; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCGGGGAGCGAGTGGAGGCTGCAGCCAGCTCGTCTCGGCGCCCGCGTGGCGTGGCGA 60
Db 1 TCGGGGAGCGAGTGGAGGCTGCAGCCAGCTCGTCTCGGCGCCCGCGTGGCGTGGCGA 60
QY 61 AGCCCCCGCCCGCTTCGCGCGCTCGGAATAGAGTCCCGGAAAGTGTGGCCATTGAG 120
Db 61 AGCCCCCGCCCGCTTCGCGCGCTCGGAATAGAGTCCCGGAAAGTGTGGCCATTGAG 120
QY 121 GCCCGAAAGCGAGCGCGGAAAGAGAGAGAAATCCGAAAGCCGAGGCTCAACAGAA 180
Db 121 GCCCGAAAGCGAGCGCGGAAAGAGAGAGAAATCCGAAAGCCGAGGCTCAACAGAA 180
QY 181 GCTCTCTTTAGTTTCTCTCTCCACCGCCACCAACACCCAGCCGACCTTTGGCAGAC 240
Db 181 GCTCTCTTTAGTTTCTCTCTCCACCGCCACCAACACCCAGCCGACCTTTGGCAGAC 240
QY 241 CCCACACCCCGGAGCGAGAGGAGATCTCTGGATCAGATGATGAGGAGCAAGAGGAC 300
Db 241 CCCACACCCCGGAGCGAGAGGAGATCTCTGGATCAGATGATGAGGAGCAAGAGGAC 300
QY 301 CCTGGGACTACTGCAAGGTGGATATCATCCAGTGAAATTTGGAGACCTTCTCAATGSC 360
Db 301 CCTGGGACTACTGCAAGGTGGATATCATCCAGTGAAATTTGGAGACCTTCTCAATGSC 360
QY 361 CGGTATCATGTTATTAGAAAGCTTGGATGGGGGCACTTCTCTACTCTGTGGCTGTGCTG 420
Db 361 CGGTATCATGTTATTAGAAAGCTTGGATGGGGGCACTTCTCTACTCTGTGGCTGTGCTG 420
QY 421 GATATGAGGGGAAAGATTGTTGCAATGAAAGTTGTAAGTCCAGCAGATTATACG 480
Db 421 GATATGAGGGGAAAGATTGTTGCAATGAAAGTTGTAAGTCCAGCAGATTATACG 480
QY 481 GAGACAGCTTGGATGAAATGCTCAATGTTTCGAGAAAGTATCCAGTATCCAGTGAC 540
Db 481 GAGACAGCTTGGATGAAATGCTCAATGTTTCGAGAAAGTATCCAGTATCCAGTGAC 540
QY 541 CCAAAACAAGACATGGTGGTCCAGCTCATTTGACGACTTCAAGATTTTCAGGCAATGGG 600
Db 541 CCAAAACAAGACATGGTGGTCCAGCTCATTTGACGACTTCAAGATTTTCAGGCAATGGG 600
QY 601 ATACATGTCTGATGGTCTTTCGAAGTACTTGGCCACCATCTCTCTCAAGTGGATCAAA 660
Db 601 ATACATGTCTGATGGTCTTTCGAAGTACTTGGCCACCATCTCTCTCAAGTGGATCAAA 660
QY 661 TCCAACTATCAAGGCTCCAGTACGTTGTGTGAAGAGTATCATTTGACAGAGTCTCTCAA 720
Db 661 TCCAACTATCAAGGCTCCAGTACGTTGTGTGAAGAGTATCATTTGACAGAGTCTCTCAA 720
QY 721 GGGTTAGATTACTTACACAGTAAAGTGAAGATCATTTACTGACATAAAGCCGGAAT 780
Db 721 GGGTTAGATTACTTACACAGTAAAGTGAAGATCATTTACTGACATAAAGCCGGAAT 780
QY 781 ATCTTTGATGTGTGGATGATGCATATGTGAGAAGATGGCAGCTGAGGCCACTGAGTGG 840
Db 781 ATCTTTGATGTGTGGATGATGCATATGTGAGAAGATGGCAGCTGAGGCCACTGAGTGG 840
QY 841 CAGAAAGCAGGTTGCTCTCTCTTCAGGCTCTGAGTGTAGTACGGCTCCACAGCAGAA 900
Db 841 CAGAAAGCAGGTTGCTCTCTCTTCAGGCTCTGAGTGTAGTACGGCTCCACAGCAGAA 900

Db 841 CAGAAAGCAGGTGCTCCTCCTTTCAGGGTCTGCAAGTGAGTACGGCTCCACAGCAGAAA 900
Qy 901 CCTATAGGAAAAATATCTAAAAACAAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 960
Db 901 CCTATAGGAAAAATATCTAAAAACAAAAAGAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 960
Qy 961 GCTGAGTTATTTGAGAGCGCTGAGAGAGCTGAGAGAGTAGAAGAAATTTGGAGCGGAGAAAGG 1020
Db 961 GCTGAGTTATTTGAGAGCGCTGAGAGAGCTGAGAGAGTAGAAGAAATTTGGAGCGGAGAAAGG 1020
Qy 1021 AAAATAATAGAAAGAAACATCACTCAGCTGCAAGCTTCCAAATGACCCAGGATCGGAAATAC 1080
Db 1021 AAAATAATAGAAAGAAACATCACTCAGCTGCAAGCTTCCAAATGACCCAGGATCGGAAATAC 1080
Qy 1081 TGCCACAGAGGTGAAACTAAAAAACAACAGGATTTAGAGAGCGGCTGAGGAGAGACTGCA 1140
Db 1081 TGCCACAGAGGTGAAACTAAAAAACAACAGGATTTAGAGAGCGGCTGAGGAGAGACTGCA 1140
Qy 1141 AAGGACAAATGTTGAGAGCTGAGGACCGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1200
Db 1141 AAGGACAAATGTTGAGAGCTGAGGACCGAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 1200
Qy 1201 GAAAGAGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1201 GAAAGAGTAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
Qy 1261 TCACCTAAAAACCAATGGCCATATTTGAGAAATGGCCCATCTCACTGGAGCAGCAATGGAC 1320
Db 1261 TCACCTAAAAACCAATGGCCATATTTGAGAAATGGCCCATCTCACTGGAGCAGCAATGGAC 1320
Qy 1321 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Db 1321 GATGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
Qy 1381 AATGAGAAAGTGAATTTACACATATAGCAGCTCCATGAAACAAATTTCAATGGTGAATTTGCA 1440
Db 1381 AATGAGAAAGTGAATTTACACATATAGCAGCTCCATGAAACAAATTTCAATGGTGAATTTGCA 1440
Qy 1441 AATGAGAAAGTGAATTTACACATATAGCAGCTCCATGAAACAAATTTCAATGGTGAATTTGCA 1500
Db 1441 AATGAGAAAGTGAATTTACACATATAGCAGCTCCATGAAACAAATTTCAATGGTGAATTTGCA 1500
Qy 1501 GATGAGAAAGTGAATTTACACATATAGCAGCTCCATGAAACAAATTTCAATGGTGAATTTGCA 1560
Db 1501 GATGAGAAAGTGAATTTACACATATAGCAGCTCCATGAAACAAATTTCAATGGTGAATTTGCA 1560
Qy 1561 CAAGAGGAGAGCTCCATCCCATGACAGAGCGGTTCAGGCTCCAGCTCCAGCTCCAGCTCCAG 1620
Db 1561 CAAGAGGAGAGCTCCATCCCATGACAGAGCGGTTCAGGCTCCAGCTCCAGCTCCAGCTCCAG 1620
Qy 1621 GATTTGCCAAAAGCAAAAAACCGGGCAGCTGACTTTGTTGGTGAATTTCCCTGGATCCGCGG 1680
Db 1621 GATTTGCCAAAAGCAAAAAACCGGGCAGCTGACTTTGTTGGTGAATTTCCCTGGATCCGCGG 1680
Qy 1681 AATGAGAAAGTGAATTTACACATATAGCAGCTCCATGAAACAAATTTCAATGGTGAATTTGCA 1740
Db 1681 AATGAGAAAGTGAATTTACACATATAGCAGCTCCATGAAACAAATTTCAATGGTGAATTTGCA 1740
Qy 1741 CACTTCACGGAAGACATCCAGAGCGGTTCAGTACCGCTCCATAGAGGTTTTTAATAGAGCG 1800
Db 1741 CACTTCACGGAAGACATCCAGAGCGGTTCAGTACCGCTCCATAGAGGTTTTTAATAGAGCG 1800
Qy 1801 GGGTACAGACCCCTCGGACATCTGGAGCAGCGGTGATGGCAATTTGAGCTGGCAAG 1860
Db 1801 GGGTACAGACCCCTCGGACATCTGGAGCAGCGGTGATGGCAATTTGAGCTGGCAAG 1860
Qy 1861 GGAGATTTATTTGAAACCAATTTGAGAAAGTGAATTTGAGAAAGTGAATTTGAGAAAGTGAAT 1920
Db 1861 GGAGATTTATTTGAAACCAATTTGAGAAAGTGAATTTGAGAAAGTGAATTTGAGAAAGTGAAT 1920
Qy 1921 GCCCACATAGAGCTGCTGAGAGTATTTCCAGGCACTTTGCTCTATCTGAGAAATAT 1980
Db 1921 GCCCACATAGAGCTGCTGAGAGTATTTCCAGGCACTTTGCTCTATCTGAGAAATAT 1980

Qy 1981 TCTCGGAAATCTTCAATCGCAGAGGAACTGGACACATCACCAAGCTGAAGCCCTGG 2040
Db 1981 TCTCGGAAATCTTCAATCGCAGAGGAACTGGACACATCACCAAGCTGAAGCCCTGG 2040
Qy 2041 AGCCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 2041 AGCCTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Qy 2101 ACAGATTTCTCGATCCCGATGTTAGAAATGGTTCAGAAAAACGAGCCCTCAGCTGGCGAA 2160
Db 2101 ACAGATTTCTCGATCCCGATGTTAGAAATGGTTCAGAAAAACGAGCCCTCAGCTGGCGAA 2160
Qy 2161 TGCCCTCGGCAATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Db 2161 TGCCCTCGGCAATCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
Qy 2221 GCAATGTTCCAGTACATTTGAGCACTTAACCGTGACTCTCATTTCTTTAACAGGATTA 2280
Db 2221 GCAATGTTCCAGTACATTTGAGCACTTAACCGTGACTCTCATTTCTTTAACAGGATTA 2280
Qy 2281 GTGAGCTGGCTTCATCTCAGACCTTTATTTGCTTTGAGGTACTGTTGTTGACATTTT 2340
Db 2281 GTGAGCTGGCTTCATCTCAGACCTTTATTTGCTTTGAGGTACTGTTGTTGACATTTT 2340
Qy 2341 GCTTTTGTGCACTGTGATCCTGGGAAAGGGTAGTCTTTTGTCTTCACTAAGTATTTA 2400
Db 2341 GCTTTTGTGCACTGTGATCCTGGGAAAGGGTAGTCTTTTGTCTTCACTAAGTATTTA 2400
Qy 2401 CTGACATTTTCTTCTGAAAAAATAACATGCTCTAAGCATGTTCTTGTGTTGTG 2460
Db 2401 CTGACATTTTCTTCTGAAAAAATAACATGCTCTAAGCATGTTCTTGTGTTGTG 2460
Qy 2461 ACATTTCAATGTCTTTTGTGAAATGAAAAATACCTTCCCTTTGTTGTTGCGAGTTT 2520
Db 2461 ACATTTCAATGTCTTTTGTGAAATGAAAAATACCTTCCCTTTGTTGTTGCGAGTTT 2520
Qy 2521 TGTAATAATTTATGAAGAAATATTTTAGTCTGAGTACTATAATTTTCAATCTTTAAGAAA 2580
Db 2521 TGTAATAATTTATGAAGAAATATTTTAGTCTGAGTACTATAATTTTCAATCTTTAAGAAA 2580
Qy 2581 TTATCAAGTTGGGAAACCAAGAAATAGCAAGGAAATGTAATTTTATCTTGGGAAA 2640
Db 2581 TTATCAAGTTGGGAAACCAAGAAATAGCAAGGAAATGTAATTTTATCTTGGGAAA 2640
Qy 2641 GGGACATCATTTCTGTTATTTATAGTATGTAATGTAATGTAATGTAATGTAATGTAATGTA 2700
Db 2641 GGGACATCATTTCTGTTATTTATAGTATGTAATGTAATGTAATGTAATGTAATGTAATGTA 2700
Qy 2701 AAATATGGAGGGGGGACTCAAAATTTCAAGAAAGCTAAAAAAGTAAATAGTAAAGG 2760
Db 2701 AAATATGGAGGGGGGACTCAAAATTTCAAGAAAGCTAAAAAAGTAAATAGTAAAGG 2760
Qy 2761 AAAATATCTTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2820
Db 2761 AAAATATCTTTATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTTAATTT 2820
Qy 2821 ATTAATACACTTAACCTTTGAGAAAGCTATGAAAAATTTAATACCATGAAAAAGATCA 2880
Db 2821 ATTAATACACTTAACCTTTGAGAAAGCTATGAAAAATTTAATACCATGAAAAAGATCA 2880
Qy 2881 AAAATCAATAATCAAAACCCCACTATAGTCCAAAGCAATTTCTCGGCGGTCAACTT 2940
Db 2881 AAAATCAATAATCAAAACCCCACTATAGTCCAAAGCAATTTCTCGGCGGTCAACTT 2940
Qy 2941 TTTAATCATTTATCTAGTACTGAGCTGGTCTCAATTTAATTTAATTTAATTTAATTTAATTT 3000
Db 2941 TTTAATCATTTATCTAGTACTGAGCTGGTCTCAATTTAATTTAATTTAATTTAATTTAATTT 3000
Qy 3001 CACCAAGTTAGGCCCGGTAATGTCTCTCTCGTGAATCTGTCATACAAATACATTTTCT 3060
Db 3001 CACCAAGTTAGGCCCGGTAATGTCTCTCTCGTGAATCTGTCATACAAATACATTTTCT 3060

QY 3061 ATTATTTAGTGGTCTCGTTTATCTTTCGCCACATCTTTGTTCACTATTTTCTAGTTA 3120
|||||
Db 3061 ATTATTTAGTGGTCTCGTTTATCTTTCGCCACATCTTTGTTCACTATTTTCTAGTTA 3120
QY 3121 CTCCTATCTTTGGCTGATTAACTCTCTCATPATACTCATATAAACTTCTGAAATTTTC 3180
Db 3121 CTCCTATCTTTGGCTGATTAACTCTCTCATPATACTCATATAAACTTCTGAAATTTTC 3180
QY 3181 ACATAAACTACTAGAGTACCTCACCATCTCTGTTTTTAAACCGGAGGAGTTACTATATA 3240
|||||
Db 3181 ACATAAACTACTAGAGTACCTCACCATCTCTGTTTTTAAACCGGAGGAGTTACTATATA 3240
QY 3241 ATTACTATTTAAA 3253
Db 3241 ATTACTATTTAAA 3253

RESULT 3

AR077238 LOCUS AR077238 2791 bp DNA linear PAT 31-AUG-2000
DEFINITION Sequence 8 from patent US 5962232.

ACCESSION AR077238

VERSION AR077238.1 GI:10003984

KEYWORDS

SOURCE Unknown.

ORGANISM

Unclassified.

REFERENCE 1 (bases 1 to 2791)

AUTHORS Bandman, O., Hillman, J. L., Lal, P., Akerblom, I. E., Shah, P.,

Corley, N. C. and Guegler, K. J.

TITLE Protein kinase molecules

JOURNAL Patent: US 5962232-A 8 05-OCT-1999;

FEATURES

source Location/Qualifiers

1..2791

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 79.8%; Score 2595.6; DB 6; Length 2791;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 2713; Conservative 0; Mismatches 14; Indels 69; Gaps 4;

QY 15 GGAGGCTGAGCCCGAGCTCGTCTCGGCGCGCGTCCGCGGAGGCGCCCGCCCG 74

Db 1 GGAGGCTGAGCCCGAGCTCGTCTCGGCGCGCGTCCGCGGAGGCGCCCGCCCG 59

QY 75 CTTCCGCGCGTCCGGAATGAGCTCCCGGAAAGTGTGGCCATTCAGGCCGGAAGCGGAG 134

Db 60 CTTCCGCGCGTCCGGAATGAGCTCCCGGAAAGTGTGGCCATTCAGGCCGGAAGCGGAG 119

QY 135 GCCGAAAGAGAGAAACATCCGAAAAA----- 161

Db 120 GCCGAAAGAGAGAAACATCCGAAAAA----- 179

QY 162 -----GCCGAGGCTCAACAGAAAGCTCTTTAG 190

Db 180 TTAACCTCTGAGAAGTCTCTTTCAGAAAGGCGGAGGCTCAACAGAAAGCTCTTTAG 239

QY 191 TTCCTCTCTCCACCGCCACACACACCGCCACCTTTGCCAGACCCACACCC 250

Db 240 TTCCTCTCTCCACCGCCACACACACCGCCACCTTTGCCAGACCCACACCC 299

QY 251 CGGAGCCAGGAGGAGATCTTGGGATCAGATGATGAGGAGCAAGGAGCCCTGCGGACT 310

Db 300 CGGAGCCAGGAGGAGATCTTGGGATCAGATGATGAGGAGCAAGGAGCCCTGCGGACT 359

QY 311 ACTGCAAAAGGTGGATATCATCCAGTGAATAATTGGAGACCTCTTCAATGGCCGGTATCATG 370

Db 360 ACTGCAAAAGGTGGATATCATCCAGTGAATAATTGGAGACCTCTTCAATGGCCGGTATCATG 419

QY 371 TTATTAGAAAGCTTGGATGGGGCACTTCTACTGTCTGGCTGTCTGGATATGCAGG 430

Db 420 TTATTAGAAAGCTTGGATGGGGCACTTCTACTGTCTGGCTGTCTGGATATGCAGG 479

QY 431 GGAAAAGATTCTTGCATGAAGTTGTAAAAAGTCCCGACGATTTATACGGAGACAGCCT 490
Db 480 GGAAAAGATTCTTGCATGAAGTTGTAAAAAGTCCCGACGATTTATACGGAGACAGCCT 539
QY 491 TCGATGAATAAAATTTGCTCAAAATGTGTTTCGAGAAAGTGTATCCAGTGCAGCCCAACAAAG 550
Db 540 TCGATGAATAAAATTTGCTCAAAATGTGTTTCGAGAAAGTGTATCCAGTGCAGCCCAACAAAG 599
QY 551 ACATGGTGTCTCAGCTCATTTGACGACTTCAAGATTTACGGATGAATGGATACATGTCT 610
Db 600 ACATGGTGTCTCAGCTCATTTGACGACTTCAAGATTTACGGATGAATGGATACATGTCT 659
QY 611 GCATGGTCTTCGAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAATCCAACTATC 670
Db 660 GCATGGTCTTCGAAGTACTTGGCCACCATCTCTCAAGTGGATCATCAATCCAACTATC 719
QY 671 AAGGCTCTCCAGTACTGTTGTGAAGAGTATCATTTGCAGAGGTCCTTTCAAGGGTTAGATT 730
Db 720 AAGGCTCTCCAGTACTGTTGTGAAGAGTATCATTTGCAGAGGTCCTTTCAAGGGTTAGATT 779
QY 731 ACTTACACAGTAAAGTCAAGATCATTTACATGACATAAAGCCGGAATAATATCTTGATGT 790
Db 780 ACTTACACAGTAAAGTCAAGATCATTTACATGACATAAAGCCGGAATAATATCTTGATGT 839
QY 791 GTGTGATGATCATATGTGAGAAATGCGCAGCTGAGGCCACTGAGTGGCAAGAGCAG 850
Db 840 GTGTGATGATCATATGTGAGAAATGCGCAGCTGAGGCCACTGAGTGGCAAGAGCAG 899
QY 851 GTGCTCTCTCTCTCAGGCTCTGAGTGTGAGTACGCTCCACAGCAGAAACCTTATAGGAA 910
Db 900 GTGCTCTCTCTCTCAGGCTCTGAGTGTGAGTACGCTCCACAGCAGAAACCTTATAGGAA 959
QY 911 AAATATCTAAAAACAAAAAGAAAAAATCTGAAAAAGAAACAGAGAGGCGAGGCTGAGTTAT 970
Db 960 AAATATCTAAAAACAAAAAGAAAAAATCTGAAAAAGAAACAGAGAGGCGAGGCTGAGTTAT 1019
QY 971 TGGAGAGCGCTCGCAGGAGATAGAAATTTGGAGGAGAGCTGAAAGAAAAATATATAG 1030
Db 1020 TGGAGAGCGCTCGCAGGAGATAGAAATTTGGAGGAGAGCTGAAAGAAAAATATATAG 1079
QY 1031 AAGAAAAACATCACTCAGCTGACCTTCCAATGACAGGATGCGCAATATCTGCCAGAGG 1090
Db 1080 AAGAAAAACATCACTCAGCTGACCTTCCAATGACAGGATGCGCAATATCTGCCAGAGG 1139
QY 1091 TGAACCTAAAAACACAGGATTTAGAGGAGCGGCTGAGCAGAGACTGCAAGAGCAATG 1150
Db 1140 TGAACCTAAAAACACAGGATTTAGAGGAGCGGCTGAGCAGAGACTGCAAGAGCAATG 1199
QY 1151 GTGAAGCTGAGGAGCAGGAGAGAAAGATGCTGAGAAAGAAACATTTGAAAAAGATG 1210
Db 1200 GTGAAGCTGAGGAGCAGGAGAGAAAGATGCTGAGAAAGAAACATTTGAAAAAGATG 1259
QY 1211 AAGATGATGTAGATCAGGAACCTTTCGAAACATAGACCTTACGTGGATAGATCACCTAAAA 1270
Db 1260 AAGATGATGTAGATCAGGAACCTTTCGAAACATAGACCTTACGTGGATAGATCACCTAAAA 1319
QY 1271 CCAATGGCCATATTGAGAAATGGCCCATTTCTACTGGAGCAGCAACTGGACGATGAAGATG 1330
Db 1320 CCAATGGCCATATTGAGAAATGGCCCATTTCTACTGGAGCAGCAACTGGACGATGAAGATG 1379
QY 1331 ATGATGAAGAGAGCTGCCCAAAATCTGTAGGAATATAATCTTGATGAGCCAAATGCAAGAA 1390
Db 1380 ATGATGAAGAGAGCTGCCCAAAATCTGTAGGAATATAATCTTGATGAGCCAAATGCAAGAA 1439
QY 1391 GTGATTTACATATAGCAGCTCTTACCAATTTCAATGGTGAATTTGCCAAATGGACGAC 1450
Db 1440 GTGATTTACATATAGCAGCTCTTACCAATTTCAATGGTGAATTTGCCAAATGGACGAC 1499
QY 1451 ATAAAAATCCCGAGTACAGTTTCCAGAGTTTTTCCACCTCGTTGTTCTCTGGATCCTTAG 1510
Db 1500 ATAAAAATCCCGAGTACAGTTTCCAGAGTTTTTCCACCTCGTTGTTCTCTGGATCCTTAG 1559

1511 AACCTGTGGCTCGGCTCTGTCTTCTGAGGATACACCTTACTGAGCAAGAGGAGA 1570
1560 AACCTGTGGCTCGGCTCTGTCTTCTGAGGATACACCTTACTGAGCAAGAGGAGA 1619
1571 GCAGTCCATCCCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1630
1620 GCAGTCCATCCCATGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1679
1631 AGCAAAACCCGCGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1690
1680 AAGCAAAACCCGCGAGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1739
1691 AATTTAGAGTAAATTTGCTGACCTGGGAAATGCTTGTGTGTGTGTGTGTGTGTGT 1750
1740 AATTTAGAGTAAATTTGCTGACCTGGGAAATGCTTGTGTGTGTGTGTGTGTGTGT 1799
1751 AAGACATCCAGAGCGGTGATACCGCTCATAGAGGTTTAAATAGAGCGGGGTACAGCA 1810
1800 AAGACATCCAGAGCGGTGATACCGCTCATAGAGGTTTAAATAGAGCGGGGTACAGCA 1859
1811 CCCCTCGGACATCTGGAGCAGCGGTGTATGGCATTTGAGCTGGCAACGGGAGATTAT 1870
1860 CCCCTCGGACATCTGGAGCAGCGGTGTATGGCATTTGAGCTGGCAACGGGAGATTAT 1919
1871 TGTTCGAACCATCTCTGGGAGAGACTATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 1930
1920 TGTTCGAACCATCTCTGGGAGAGACTATTCAGAGAGAGAGAGAGAGAGAGAGAGAG 1979
1931 TAGAGTGTAGGAGTATTCAGAGCAGCTTTGCTCTATCTGGAATATTTCTCGGGAT 1990
1980 TAGAGTGTAGGAGTATTCAGAGCAGCTTTGCTCTATCTGGAATATTTCTCGGGAT 2039
1991 TCTTCATCGCAG 2050
2040 TCTTCATCGCAG 2099
2051 ATGACTCTGTGGAAGAGTATGGCTGGCCCATGAAGATCTGACAGAGTTTACAGATTCC 2110
2100 ATGACTCTGTGGAAGAGTATGGCTGGCCCATGAAGATCTGACAGAGTTTACAGATTCC 2159
2111 TGATCCGATTTAGAAATGTTTCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2170
2160 TGATCCGATTTAGAAATGTTTCCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2219
2171 ATCTCTGTGTGAATCTTAGCAATTTCTACCAATTTGATCTGAGCTAGCAATTTGTC 2230
2220 ATCTCTGTGTGAATCTTAGCAATTTCTACCAATTTGATCTGAGCTAGCAATTTGTC 2279
2231 CCAGTACATTTGACCTAAACGGTGACTCTCAATCTTTTACAGAGATTACAGTGGCTGC 2290
2280 CCAGTACATTTGACCTAAACGGTGACTCTCAATCTTTTACAGAGATTACAGTGGCTGC 2339
2291 TTCACTCTGAGCTTTATTTCTGTTGAGTACTGTTGTTTACAGATTGCTTTTGTG 2350
2340 TTCACTCTGAGCTTTATTTCTGTTGAGTACTGTTGTTTACAGATTGCTTTTGTG 2399
2351 CACTGTGATCTGCGGAGAGGTAGTCTTTTGTCTTCTGAGTAAAGTATTACTGACCAATT 2410
2400 CACTGTGATCTGCGGAGAGGTAGTCTTTTGTCTTCTGAGTAAAGTATTACTGACCAATT 2459
2411 TCTTCTGGAACCAATATCATGTCTTAAGCATTTGTTTGTGTTGTTGTTGTTGTTGTTG 2470
2460 TCTTCTGGAACCAATATCATGTCTTAAGCATTTGTTTGTGTTGTTGTTGTTGTTGTTG 2519
2471 GTCAATTTTTTGAATGAAAATACCTTTCCCTTGTGTTTGGCAGGTTTGTGAATATT 2530
2520 GTCAATTTTTTGAATGAAAATACCTTTCCCTTGTGTTTGGCAGGTTTGTGAATATT 2579
2531 TATGAAGAAATATTTTGTAGCTGAGTACTATATATATATATATATATATATATATATAT 2590
2580 TATGAAGAAATATTTTGTAGCTGAGTACTATATATATATATATATATATATATATATAT 2639
2591 GGGAACCAAGAAATAGCAGGGAATGTACAAATTTTATCTTCTGGCAAGGAGCATCAT 2650

2640 GGAACCAAGAAATAGCAGGGAATGTACAAATTTTATCTTCTGCAAGGAGCATCAT 2696
2651 TCTGTATTATAGTGTATGTAAATGCACCTGTAAATGTACTTTGGATTAATATGGA 2710
2697 TCTGTATTATAGTGTATGTAAATGCACCTGTAAATGTACTTTCCATTAAATATGGA 2756
2711 GGGGGAGCTCAATTTTCAGAAAAGCTTAAAAAAA 2746
2757 -GGGGAGCTCAATTTTCAGAAAAGCTTAAAAAAA 2791

RESULT 4
BD107902
LOCUS
DEFINITION Protein kinase molecules.
ACCESSION BD107902
VERSION BD107902.1 GI:23202720
KEYWORDS JP 2002501749-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2791)
AUTHORS Bandman, O., Hillman, J.L., Lai, P., Akerblom, I.E., Shah, P.,
Corley, N.C. and Guegler, K.J.
TITLE Protein kinase molecules
JOURNAL Patent: JP 2002501749-A 2 22-JAN-2002;
INCYTE PHARMACEUTICALS INC
OS Homo sapiens (human)
PN JP 2002501749-A/2
PD 22-JAN-2002
PF 12-JAN-1999 JP 2000529440
PR 30-JAN-1998 US 09/016000
PI OLGA BANDMAN, JENNIFER L HILLMAN, PREETI LAL, INGRID E AKERBLOM,
PI PURVI SHAH,
PI NEIL C CORLEY, KARL J GUEGLER
PC C12N15/09,A61K38/45,A61K45/00,C07K16/40,C12N1/15,C12N1/19, PC
C12N1/21,
PC C12N5/10,C12N9/12,C12Q1/68,G01N33/68,C12N15/00,A61K37/52,C12N5/ PC
00
CC 307624 Location/Qualifiers
FH Key 1..2791 /organism='Homo sapiens (human)'.
FT source Location/Qualifiers
FT 1..2791 /organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

FEATURES
source

ORIGIN
Query Match 79.8%; Score 2595.6; DB 6; Length 2791;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2713; Conservative 0; Mismatches 14; Indels 69; Gaps 4;
QY 15 GGAGGCTGAGCCAGCTCTGCTCGGCCCGCGCTCGCGCGCGCGCGCGCGCGCGCGCG 74
DB 1 GGAGGCTGAGCCAGCTCTGCTCGGCCCGCGCTCGCTCGCGCGCGCGCGCGCGCGCG 59
QY 75 CTTCCGCGCGCTCGGAATGAGCTCCCGGAAGTGTGCTGCGCATTCAGGCGCCGAAAGCGGAG 134
DB 60 CTTCCGCGCGCTCGGAATGAGCTCCCGGAAGTGTGCTGCGCATTCAGGCGCCGAAAGCGGAG 119
QY 135 GCCGAAAAGAGAAACATCCGAAAAA----- 161
DB 120 GCCGAAAAGAGAAACATCCGAAAAA-----GCCGAGCCTCAACAGAAAGCTCTCTTAG 179
QY 162 -----GCCGAGCCTCAACAGAAAGCTCTCTTAG 190
DB 180 TTAACCTGAGAAAGTCTCTCTTCAGAAAGCGCGCGCTCAACAGAAAGCTCTCTTAG 239

QY 393 GCACCTTCTCTACTGTCTGCTGCTGGGATATGCAAGGGGAAAAGATTTGTTGCAATGAA 452
DB 423 GCACCTTCTCTACTGTCTGCTGCTGGGATATGCAAGGGGAAAAGATTTGTTGCAATGAA 482
QY 453 AGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAAATAAAATTTGCTCAA 512
DB 483 AGTTGTAAAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAAATAAAATTTGCTCAA 542
QY 513 ATGTGTTTCAGAAAGTGATCCCAAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATTTGA 572
DB 543 ATGTGTTTCAGAAAGTGATCCCAAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATTTGA 602
QY 573 CGACCTTCAAGATTTTCAGGCAATGAAATGGGATACATGTCTGATGGTCTTCGAAAGTACTTGG 632
DB 603 CGACCTTCAAGATTTTCAGGCAATGAAATGGGATACATGTCTGATGGTCTTCGAAAGTACTTGG 662
QY 633 CCACCATCTCTCAAGTGGATCATCAAAATCCAACATATCAAGGCCTCCCAAGTACGTTGTGT 692
DB 663 CCACCATCTCTCAAGTGGATCATCAAAATCCAACATATCAAGGCCTCCCAAGTACGTTGTGT 722
QY 693 GAAGAGTATCATTTCCAGAGTCTCTCAAGGGTTAGATTACTTACACAGTAAGTGCAGAT 752
DB 723 GAAGAGTATCATTTCCAGAGTCTCTCAAGGGTTAGATTACTTACACAGTAAGTGCAGAT 782
QY 753 CATTTCACTGACATAAAGCCGGAATAATCTTGTGCTGTGGATGATGTCATATGTGAG 812
DB 783 CATTTCACTGACATAAAGCCGGAATAATCTTGTGCTGTGGATGATGTCATATGTGAG 842
QY 813 AAGAAATGGCAGCTGAGGCCACCTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCAAGGGTC 872
DB 843 AAGAAATGGCAGCTGAGGCCACCTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCAAGGGTC 902
QY 873 TGCAGTGAGTACGGCTCCACAGCAGAAACCTATAGGAAATAATCTTAAATAACAAAGAA 932
DB 903 TGCAGTGAGTACGGCTCCACAGCAGAAACCTATAGGAAATAATCTTAAATAACAAAGAA 962
QY 933 AAAAAGTGAAGAAAGAACAGAGGCGGCTGAGTTATTTGGAGAGCGGCTTCAGGAGAT 992
DB 963 AAAAAGTGAAGAAAGAACAGAGGCGGCTGAGTTATTTGGAGAGCGGCTTCAGGAGAT 1022
QY 993 AGAAGAAATGGAGCGAGAGCTGAAAGGAAATAATAGAAAGAAACATCACCTCAGCTGC 1052
DB 1023 AGAAGAAATGGAGCGAGAGCTGAAAGGAAATAATAGAAAGAAACATCACCTCAGCTGC 1082
QY 1053 ACCTTCCAAATGACAGGATGGCGAATATCTGCCAGAGGTGAAATCTTAAATAACAAAGAA 1112
DB 1083 ACCTTCCAAATGACAGGATGGCGAATATCTGCCAGAGGTGAAATCTTAAATAACAAAGAA 1142
QY 1113 AGAGAGGCGGCTGAGGCGAGAGCTGCAAGAGCAATGGTGAAGCTGAGGACAGGAGAA 1172
DB 1143 AGAGAGGCGGCTGAGGCGAGAGCTGCAAGAGCAATGGTGAAGCTGAGGACAGGAGAA 1202
QY 1173 GAAAGAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGTAGATCAGGAATCT 1232
DB 1203 GAAAGAGATGCTGAGAAAGAAACATTTGAAAGAGATGAAGATGATGTAGATCAGGAATCT 1262
QY 1233 TCGCAACATPAGACCTTACGTGGATGAAATCACCTAAACCAATGGCATAATTTGAGAAATGG 1292
DB 1263 TCGCAACATPAGACCTTACGTGGATGAAATCACCTAAACCAATGGCATAATTTGAGAAATGG 1322
QY 1293 CCCATTCTCAGTGGAGCAGCAACTGGAACGATGAAGATGATGAAGAGAGACTGCCCCAAA 1352
DB 1323 CCCATTCTCAGTGGAGCAGCAACTGGAACGATGAAGATGATGAAGAGAGACTGCCCCAAA 1382
QY 1353 TCCTGAGGAATAATCTTTGATGAGCCAAATCGAAGAGTGAATACACATATAGCAGCTC 1412
DB 1383 TCCTGAGGAATAATCTTTGATGAGCCAAATCGAAGAGTGAATACACATATAGCAGCTC 1442
QY 1413 CTATGAACAAATTTCAATGGTGAATTTGCCAAATGGAACGACATAAAATTTCCGAGTCAAGTT 1472
DB 1443 CTATGAACAAATTTCAATGGTGAATTTGCCAAATGGAACGACATAAAATTTCCGAGTCAAGTT 1502

QY 1473 CCCAGAGTTTTCCACTCGTGTCTCTCTGATCCCTTAGAACCTTGTGGCCTCGGCTCTGT 1532
DB 1503 CCCAGAGTTTTCCACTCGTGTCTCTCTGATCCCTTAGAACCTTGTGGCCTCGGCTCTGT 1562
QY 1533 GCTTTCTGAGGATCACCACTTTACTGAGCAAGAGGAGAGCAGTCCATCCCATGACAGAAG 1592
DB 1563 GCTTTCTGAGGATCACCACTTTACTGAGCAAGAGGAGAGCAGTCCATCCCATGACAGAAG 1622
QY 1593 CAGAACGGTTTCAGGCTCAGTACTGCGGATTTGCCAAAGCAAAACCCGGGCGCTGGA 1652
DB 1623 CAGAACGGTTTCAGGCTCAGTACTGCGGATTTGCCAAAGCAAAACCCGGGCGCTGGA 1682
QY 1653 CTTGTTGGTGAATCCCTGAGTCCCGGAAATCGAGATAAAATTTAGAGTAAAAATTTCTCTGA 1712
DB 1683 CTTGTTGGTGAATCCCTGAGTCCCGGAAATCGAGATAAAATTTAGAGTAAAAATTTCTCTGA 1742
QY 1713 CCTGGGAAATGCTTGTGGGTGCATAAAACATTTACGGAAGACATCCAGACGCGTCAGTA 1772
DB 1743 CCTGGGAAATGCTTGTGGGTGCATAAAACATTTACGGAAGACATCCAGACGCGTCAGTA 1802
QY 1773 CCGCTCCATAGAGGTTTTTAATAGGAGCGGGGTACAGACCCCTTGGGACATCTGGAGCAC 1832
DB 1803 CCGCTCCATAGAGGTTTTTAATAGGAGCGGGGTACAGACCCCTTGGGACATCTGGAGCAC 1862
QY 1833 GSCGTGTATGGCATTTGAGCTGSCAAACGGGAGATTTATTTGTTTGAACCAATTTCTGGGA 1892
DB 1863 GSCGTGTATGGCATTTGAGCTGSCAAACGGGAGATTTATTTGTTTGAACCAATTTCTGGGA 1922
QY 1893 AGACTATTCCAGAGACGAAGACACATAGCCCCACATCATAGAGCTGCTAGGAGCATTTCC 1952
DB 1923 AGACTATTCCAGAGACGAAGACACATAGCCCCACATCATAGAGCTGCTAGGAGCATTTCC 1982
QY 1953 AAGGCATTTTGTCTATCTGGAAATAATTTCTCGGAAATTTCTTCAATCGCAGAGGAGAACT 2012
DB 1983 AAGGCATTTTGTCTATCTGGAAATAATTTCTCGGAAATTTCTTCAATCGCAGAGGAGAACT 2042
QY 2013 GGCACACATCACCAAGCTGAAAGCCCTGGAGCCTCTTGTGATGACTTGTGGAAGATATGG 2072
DB 2043 GGCACACATCACCAAGCTGAAAGCCCTGGAGCCTCTTGTGATGACTTGTGGAAGATATGG 2102
QY 2073 CTGGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTCTGATCCCGATCTTGAAGATGGT 2132
DB 2103 CTGGGCCCATGAAGATGCTGCACAGTTTACAGATTTCTCTGATCCCGATCTTGAAGATGGT 2162
QY 2133 TCCAGAAAAACGAGCCTCAGCTGGCGAATGCTTGGCAATCTTGGTGAATTTCTTGAACA 2192
DB 2163 TCCAGAAAAACGAGCCTCAGCTGGCGAATGCTTGGCAATCTTGGTGAATTTCTTGAACA 2222
QY 2193 AATTTCTACCAATATTGCAATTTGAGCTAGCAAAATGTTCCAGTACATTTGGACCTAAACGG 2252
DB 2223 AATTTCTACCAATATTGCAATTTGAGCTAGCAAAATGTTCCAGTACATTTGGACCTAAACGG 2282
QY 2253 TGACTCTCATTTCTTAAACAGGATTAACAGTGAAGTGGCTTTCATCTCAGACCTTTATTTT 2312
DB 2283 TGACTCTCATTTCTTAAACAGGATTAACAGTGAAGTGGCTTTCATCTCAGACCTTTATTTT 2342
QY 2313 GCTTTGAGGTACTGTGTGTTGACATTTTGTCTTTTGTGCACTGTGATCTCTGGGGAAGGT 2372
DB 2343 GCTTTGAGGTACTGTGTGTTGACATTTTGTCTTTTGTGCACTGTGATCTCTGGGGAAGGT 2402
QY 2373 AGTCTTTTGTCTTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2432
DB 2403 AGTCTTTTGTCTTCAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 2462
QY 2433 CTCTAAGCAATTTGTTTCTTGTGTGTGACATTTCAAAATGTCATTTTGTGAAATGAAAAAT 2492
DB 2463 CTCTAAGCAATTTGTTTCTTGTGTGTGACATTTCAAAATGTCATTTTGTGAAATGAAAAAT 2522
QY 2493 ACTTTTCCCTTTGTGTTTGGCAGGTTTGTAACTATTTATGAAGAAAAATTTTGTAGCTGA 2552
DB 2523 ACTTTTCCCTTTGTGTTTGGCAGGTTTGTAACTATTTATGAAGAAAAATTTTGTAGCTGA 2582
QY 2553 GTACTATATATTTTACAATCTTTAAGAAATTTATCAAGTGGGGAACCAAGAAAAATAGCAAG 2612

Db	2583	GTACTATATAATTTACAATCTTAAGAAATTTATCAAGTTGGAACCAAGAAATA---GCAAG	2639
Qy	2613	GAATGTGACAAATTTATCTTCTGCGAAGGACATCATTTCTGTATATATAGTGTATGTAA	2672
Db	2640	GAATGTGACAAATTTATCTTCTGCGAAGGACATCATTTCTGTATATATAGTGTATGTAA	2699
Qy	2673	ATGCACCTCTGAAATGTTACTTTTGGATTAAATATGGAGGGGACCTCAAAATTTACAGAAA	2732
Db	2700	ATGCACCTCTGAAATGTTACTTTTCCATTAATATATGGA-GGGGACTCAAAATTTACAGAAA	2758
Qy	2733	AGCTAAAAA	2741
Db	2759	AGCTACCAA	2767
RESULT 6			
AX429244			
LOCUS	AX429244	3745 bp	DNA linear PAT 21-JUN-2002
DEFINITION	Sequence 21 from Patent EP1201765.		
ACCESSION	AX429244		
VERSION	AX429244.1	GI:21540552	
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	artificial sequences.		
REFERENCE	1		
AUTHORS	Schubart,D., Habenberger,P., Stein-Gerlach,M. and Bevec,D.		
TITLE	Cellular kinases involved in cytomegalovirus infection and their inhibition		
JOURNAL	Patent: EP 1201765-A 21 02-MAY-2002;		
FEATURES	Axxima Pharmaceuticals Aktiengesellschaft (DE)		
source	Location/Qualifiers		
	1..3745		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="N/A"		
ORIGIN			
	Query Match	76.8%;	Score 2497; DB 6; Length 3745;
	Best Local Similarity	98.8%;	Pred. No. 0;
	Matches 2559; Conservative	0; Mismatches 20; Indels 10; Gaps 4;	
Qy	153	TCGGAAGAAGCGGAGCCTCAACAGAAGCTCTTTAGTTCTCTCTCCACCGCCACC	212
Db	217	TTCAGAAGGCGGAGCCTCAACAGAAGCTCTTTAGTTCTCTCTCCACCGCCACC	276
Qy	213	ACCACCAACCGCCACCTTTGCCAGACCCACACACCCCGGAGCCAGAGGAGATCCT	272
Db	277	ACCACCAACCGCCACCTTTGCCAGACCCACACACCCCGGAGCCAGAGGAGATCCT	336
Qy	273	GGGATCAGATGATGAGGAGCAGAGGACCTTGGCGGACCTGCGGACCTGCGGAGATCATCC	332
Db	337	GGGATCAGATGATGAGGAGCAGAGGACCTTGGCGGACCTGCGGAGATCATCATCC	396
Qy	333	AGTGAAAATTCGAGACCTCTTCATGGCCGGGTATCATGTTATTAGAAGCTTGGATGGG	392
Db	397	AGTGAAAATTCGAGACCTCTTCATGGCCGGGTATCATGTTATTAGAAGCTTGGATGGG	456
Qy	393	GCATCTCTACTGTCGGCTGTGCTGGGATATGCGGGGAAAAAGATTTGTTGCAATGAA	452
Db	457	GCATCTCTACTGTCGGCTGTGCTGGGATATGCGGGGAAAAAGATTTGTTGCAATGAA	516
Qy	453	AGTTGTAAAAAGTGCCCAAGCAATTATACGAGACAGCCTTGGATGAAATAAAATTTGCTCAA	512
Db	517	AGTTGTAAAAAGTGCCCAAGCAATTATACGAGACAGCCTTGGATGAAATAAAATTTGCTCAA	576
Qy	513	ATGTGTTCCGAGAAAGTGATCCAGTGACCCAAACAAACACATGGTGGTCCAGCTCATTTGA	572
Db	577	ATGTGTTCCGAGAAAGTGATCCAGTGACCCAAACAAACACATGGTGGTCCAGCTCATTTGA	636
Qy	573	CGACTTCAAGATTTTCAGGCATGAATPGGGATACATGCTGTCATGGTCTTCGAAAGTACTTGG	632

Db 1714 CTTGTTGGTGAATCCCTCGGATCCGCGAATCGAGATAAAATTTAGAGTAAATTTGCTGA 1773

Qy 1713 CTTGGGAATGCTTTGGTGCATAAAACATTTCCAGGAAGACATCCAGACGGCTCAGTA 1772

Db 1774 CTTGGGAATGCTTTGGTGCATAAAACATTTCCAGGAAGACATCCAGACGGCTCAGTA 1833

Qy 1773 CCCTCCATAGAGGTTTAAATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAGCAC 1832

Db 1834 CCCTCCATAGAGGTTTAAATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAGCAC 1893

Qy 1833 GCGGTGATAGGCAATTTGAGCTGGCAACCGGAGATTATTTGTTTGAACCAATTTCTGGGA 1892

Db 1894 GCGGTGATAGGCAATTTGAGCTGGCAACCGGAGATTATTTGTTTGAACCAATTTCTGGGA 1953

Qy 1893 AGACTATTCAGAGACGAAGACACATAGCCACATCATAGAGCTGCTAGGAGTATTC 1952

Db 1954 AGACTATTCAGAGACGAAGACACATAGCCACATCATAGAGCTGCTAGGAGTATTC 2013

Qy 1953 AAGGCATTTGCTCTATCTGGAATAATTTCTGGGAATTTCTCAATCGCAGAGGAACT 2012

Db 2014 AAGGCATTTGCTCTATCTGGAATAATTTCTGGGAATTTCTCAATCGCAGAGGAACT 2073

Qy 2013 GCGACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTAATTTGTGAAAAGTATGG 2072

Db 2074 GCGACATCACCAAGCTGAAGCCCTGGAGCCTCTTTGATGTAATTTGTGAAAAGTATGG 2133

Qy 2073 CTGGCCCATGAAATGCTGACAGTTTACAGATTTCTCGATCCCGATGTTAGAAATGGT 2132

Db 2134 CTGGCCCATGAAATGCTGACAGTTTACAGATTTCTCGATCCCGATGTTAGAAATGGT 2193

Qy 2133 TCAGAAAAACGAGCCTCAGCTGGCAATGCTTCGGCAATTTCTCAATCGCAGAGGAACT 2192

Db 2194 TCAGAAAAACGAGCCTCAGCTGGCAATGCTTCGGCAATTTCTCAATCGCAGAGGAACT 2250

Qy 2193 AATTCTACCAATTTGCAATTTGAGCTAGCAATGTTCCAGTACATTTGACCTAAACCG 2252

Db 2251 AATTCTACCAATTTGCAATTTGAGCTAGCAATGTTCCAGTACATTTGACCTAAACCG 2310

Qy 2253 TGAATCTCATTTCTTAAACAGGATTA CAAGTACCTGGCTTCACTTCCAGACCTTTATTTT 2312

Db 2311 TGAATCTCATTTCTTAAACAGGATTA CAAGTACCTGGCTTCACTTCCAGACCTTTATTTT 2370

Qy 2313 GCTTTGAGTACGTTGTTTGAATTTGCTTTTTCGTCGACCTGATCTTGGGGAAGGCT 2372

Db 2371 GCTTTGAGTACGTTGTTTGAATTTGCTTTTTCGTCGACCTGATCTTGGGGAAGGCT 2430

Qy 2373 AGTCTTTTGTCTTCAAGTAAATTTTACTGACATTTTCTTCTGGAACCAATTAACATGT 2432

Db 2431 AGTCTTTTGTCTTCAAGTAAATTTTACTGACATTTTCTTCTGGAACCAATTAACATGT 2490

Qy 2433 CTCTAAGCAATTTGTTTCTGTTGTGTGACATTTCAAAATGTCATTTTTTTTGAATGAAAT 2492

Db 2491 CTCTAAGCAATTTGTTTCTGTTGTGTGACATTTCAAAATGTCATTTTTTTTGAATGAAAT 2550

Qy 2493 ACTTTCCCTTTGTTTGGCAGGTTTGTACTATTTATGAGAAATATTTTACTGTA 2552

Db 2551 ACTTTCCCTTTGTTTGGCAGGTTTGTACTATTTATGAGAAATATTTTACTGTA 2610

Qy 2553 GTACTATAAATTTA CAATCTTAAAGAAATTTATCAAGTTGGGAACCAAGAAATAGCAAG 2612

Db 2611 GTACTATAAATTTA CAATCTTAAAGAAATTTATCAAGTTGGGAACCAAGAAATAGCAAG 2667

Qy 2613 GAAATGTACAAATTTTATCTTCTGGCAAGGACATCATTTCTGTTATTTATAGTGTATGTA 2672

Db 2668 GAAATGTACAAATTTTATCTTCTGGCAAGGACATCATTTCTGTTATTTATAGTGTATGTA 2727

Qy 2673 ATGCACCTGTAAATTTTATCTTCTGGCAAGGACATCATTTCTGTTATTTATAGTGTATGTA 2732

Db 2728 ATGCACCTGTAAATTTTATCTTCTGGCAAGGACATCATTTCTGTTATTTATAGTGTATGTA 2786

Qy 2733 AGCTAAAAA 2741

Db 2787 AGCTACCA 2795

RESULT 7
AX666550
LOCUS
DEFINITION
SEQUENCE 3 from Patent WO02094796.
ACCESSION
AX666550
VERSION
AX666550.1 GI:29291032
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
Pato, J., Keri, G., Oerfi, L., Waczek, F., Horvath, Z., Banhegyi, P., Szabadkai, I., Marosfalvi, J., Hegymegi-Barakonyi, B., sz Kelyhidi, Z., Greff, Z., Choldas, A., Bacher, G., Daub, H., Obert, S., Kurtenbach, A. and Habenberger, P.
TITLE
Benzo[g]quinoxaline derivatives as effective compounds against infectious diseases
JOURNAL
Patent: WO 02094796-A 3 28-NOV-2002;
Axxima Pharmaceuticals Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
1..3745
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="N/A"
ORIGIN
Query Match 76.8%; Score 2497; DB 6; Length 3745;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 20; Indels 10; Gaps 4;
Qy 153 TCCGAAAAAGCCGAGCCTCAACAGAAAGCTCCTTTAGTTCCTCTCCCTCCACCGCCACC 212
Db 217 TTCAGAAAGCCGAGCCTCAACAGAAAGCTCCTTTAGTTCCTCTCCCTCCACCGCCACC 276
Qy 213 ACCACCAACCGCCACCTTTGCCAGACCCCAACACCCCGAGCCAGAGAGAGATCCT 272
Db 277 ACCACCAACCGCCACCTTTGCCAGACCCCAACACCCCGAGCCAGAGAGAGATCCT 336
Qy 273 GGGATCAGATGATGAGGACAGAGAGACCTCGGACTACTGCAAGGTGGATATCATCC 332
Db 337 GGGATCAGATGATGAGGACAGAGAGACCTCGGACTACTGCAAGGTGGATATCATCC 396
Qy 333 AGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 392
Db 397 AGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 456
Qy 393 GCATTTCTCTACTGTCTGTGCTGTGGATATGCAAGGGGAAAAGATTTGTTGCAATGAA 452
Db 457 GCATTTCTCTACTGTCTGTGCTGTGGATATGCAAGGGGAAAAGATTTGTTGCAATGAA 516
Qy 453 AGTTGTAAAAGTGGCCAGCATTTATCGGAGACAGCCTTTGGATGAATAAATTTGCTCAA 512
Db 517 AGTTGTAAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAATAAATTTGCTCAA 576
Qy 513 ATGTGTTTCAGAAAGTGAATCCAGTGACCCAAACAAAGACATGCTGGTCCAGCTCATTTGA 572
Db 577 ATGTGTTTCAGAAAGTGAATCCAGTGACCCAAACAAAGACATGCTGGTCCAGCTCATTTGA 636
Qy 573 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGATGGTCTTTCGAACTACTTGG 632
Db 637 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGATGGTCTTTCGAACTACTTGG 696
Qy 633 CCACCATCTCTCAAGTGGATCATCAAATCCAACTATCAAGGCTCCAGTACTTGTGT 692
Db 697 CCACCATCTCTCAAGTGGATCATCAAATCCAACTATCAAGGCTCCAGTACTTGTGT 756
Qy 693 GAAGAGTATCATTTCCAGAGTCTCTTCAAGGGTTAGATTACTTACACAGTAAGTGAAGAT 752
Db 757 GAAGAGTATCATTTCCAGAGTCTCTTCAAGGGTTAGATTACTTACACAGTAAGTGAAGAT 816

ORGANISM	Homo sapiens
REFERENCE	Bukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Wang, H.Y., Lin, W., Dyck, J.A., Yeakley, J.M., Songyang, Z., Cantley, L.C. and Fu, X.D.
TITLE	SRPK2: a differentially expressed SR protein-specific kinase involved in mediating the interaction and localization of pre-mRNA splicing factors in mammalian cells
JOURNAL	J. Cell Biol. 140 (4), 737-750 (1998)
MEDLINE	98139536
PUBMED	9472028
REFERENCE	2 (bases 1 to 3745)
AUTHORS	Wang, H.Y., Wen, L. and Fu, X.D.
TITLE	Direct Submission
JOURNAL	Submitted (06-FEB-1997) Cell. Mol. Medicine, University of California at San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0651, USA
FEATURES	Location/Qualifiers
Source	1..3745
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
CDS	188..2248
	/note="similar to human serine kinase SRPK1, encoded by GenBank Accession Number U09564; specific for serine/arginine-rich splicing factors"
	/codon_start=1
	/product="serine kinase SRPK2"
	/protein_id="AAC05299.1"
	/db_xref="GI:1857944"
	/translation="MSVNSEKSSSRPEPQQAELVPPPPPPPPPPPLDPPTPEEILGSDDEQEDPADYCKGGYHPVKIGDLFNGRYHVRKLGHFSTWLCWDMQGRFAMVKVKSQAQHYETALDEIKLLKCRSDSPDNKMVQLIIDDFKISGNNGIHYCMVEVLGHLLKWIISKNYQGLPVRCVKSIIRVLOGLDYLSKKNKKLKKOKNLVCMVDAYVRMAEPWQKAGAPPSPGSAVTAPOQPIGKISKNKKLKKOKROALELEKLOIELEERAERKIIENITSAPSNDQDGEYCPVKLKYTGLEAAEATAKDNGAEQOEKEDEKENIEKEDDDVQELANDIDPTWIESFKTHGNEPFSLEQDDDEDDPCPEEYNLDENPAESDYTSYSEQFNGELPNRHKHPIESQFPEFSLFSGSLPACVSVLSGLTEQEESSPSHRSRTVSASSTGDLPKAKTRAA DLLVNPLOPRNDRKTRVKTADLGNACVWHKHFTEDIQTRQYRSIEVLGAGYSTPADIWSTAMAPLATGDLVFEPSHGEDYSRDEDAHIEILLGSI PRFALSGKYSRFFEN REGELRHITKLAPWSLFDVLVEKYGWPHDEAAQFTDFLIPMLEMPEKRSAGECRHP WLNS"
ORIGIN	
Query Match	76.8%; Score 2497; DB 9; Length 3745;
Best Local Similarity	98.8%; Pred. No. 0;
Matches	2559; Conservative 0; Mismatches 20; Indels 10; Gaps 4;
QY	153 TCGAAAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCTCCACCGCCACC 212
DB	217 TT'CAGAAAGCCGGAGCCTCAACAGAAAGCTCCTTTAGTTCTCTCTCTCCACCGCCACC 276
QY	213 ACCACCAACCGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAGAGAGATCCT 272
DB	277 ACCACCAACCGCCACCTTTGCCAGACCCACACCCCGGAGCCAGAGAGAGATCCT 336
QY	273 GGGATCAGATGATGAGGACAGAGGACCTCGGGACTACTGCAAGAGGTGGATATCATCC 332
DB	337 GGGATCAGATGATGAGGACAGAGGACCTCGGGACTACTGCAAGAGGTGGATATCATCC 396
QY	333 AGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 392
DB	397 AGTGAATAATGGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 456
QY	393 GCATCTCTACTGCTGCTGGATATGAGGAGGAAAAAGATTGTTGGCATGAA 452
DB	457 GCATCTCTACTGCTGCTGGATATGAGGAGGAAAAAGATTGTTGGCATGAA 516
QY	453 AGTTGTAATAAGTGCCAGCATTATACGGAGACAGCCTTTGGATGAATAAATTTGCTCAA 512
DB	517 AGTTGTAATAAGTGCCAGCATTATACGGAGACAGCCTTTGGATGAATAAATTTGCTCAA 576

QY	513 ATGTGTTTCGAGAAAGTATCCAGTGAACCCAAACAAAGACATGGTGGTCAGCTCATTTGA 572
DB	577 ATGTGTTTCGAGAAAGTATCCAGTGAACCCAAACAAAGACATGGTGGTCAGCTCATTTGA 636
QY	573 CCACCTTCAAGATTTCAGGCATGAATGGGATACATGTCATGCTCTTCCGAAGTACTTGG 632
DB	637 CCACCTTCAAGATTTCAGGCATGAATGGGATACATGTCATGCTCTTCCGAAGTACTTGG 696
QY	633 CCACCATCTCTCAAGTGGATCATCAAAATCCAACCTATCAAGGCCTCCAGTACGTTGTGT 692
DB	697 CCACCATCTCTCAAGTGGATCATCAAAATCCAACCTATCAAGGCCTCCAGTACGTTGTGT 756
QY	693 GAAGAGTATCATTCACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGAAGAT 752
DB	757 GAAGAGTATCATTCACAGGTCCTTCAAGGGTTAGATTACTTACACAGTAAGTGAAGAT 816
QY	753 CATTCATCTGACATAAAGCCGGAATAATCTTGTATGCTGTGGATGATGCATATGTGAG 812
DB	817 CATTCATCTGACATAAAGCCGGAATAATCTTGTATGCTGTGGATGATGCATATGTGAG 876
QY	813 AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTTCATCAGGTC 872
DB	877 AAGAAATGGCAGCTGAG---CCTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTTCAGGTC 933
QY	873 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTAAAAACAAAAAGAA 932
DB	934 TGCAGTGTAGTACGGCTCCACAGCAGAAACCTATAGGAAAAATATCTAAAAACAAAAAGAA 993
QY	933 AAAAATGAAAAAGAAACAGACAGCAGGCTCAGTTATTGGAGAGAGCGCTCGCAGGAGAT 992
DB	994 AAAAATGAAAAAGAAACAGACAGGAGCGGCTCAGTTATTGGAGAGAGCGCTCGCAGGAGAT 1053
QY	993 AGAAGAAATTTGGAGCGAGAGCTGAAAGGAAAAATAATAGAGAAAAACATCACCTCAGCTGC 1052
DB	1054 AGAAGAAATTTGGAGCGAGAGCTGAAAGGAAAAATAATAGAGAAAAACATCACCTCAGCTGC 1113
QY	1053 ACCTTCCAATGATCCAGGATGGCGAATACTGCCAGAGGTGAAACTTAAAAACAAACAGGATT 1112
DB	1114 ACCTTCCAATGATCCAGGATGGCGAATACTGCCAGAGGTGAAACTTAAAAACAAACAGGATT 1173
QY	1113 AGAGAGGCGGCTGAGGCGAGACACTGCAAGAGCACAATGCTGAAGCTGAGGACCCAGCAAGA 1172
DB	1174 AGAGAGGCGGCTGAGGCGAGACACTGCAAGAGCACAATGCTGAAGCTGAGGACCCAGCAAGA 1233
QY	1173 GAAAGAGATGCTGAGAAAGAAAAACATTTGAAAAAGATGAAGATGATGTAGATCAGGAAC 1232
DB	1234 GAAAGAGATGCTGAGAAAGAAAAACATTTGAAAAAGATGAAGATGATGTAGATCAGGAAC 1293
QY	1233 TCGCAACATAGACCCCTACGTGGATAGAACTCACCTTAAACCAATGGCCATATTGAGAAATGG 1292
DB	1294 TCGCAACATAGACCCCTACGTGGATAGAACTCACCTTAAACCAATGGCCATATTGAGAAATGG 1353
QY	1293 CCCATTCTCAGTGGAGCGCAACTGGACGATCAAGATGATGATGAAGAGACTGCCCAAA 1352
DB	1354 CCCATTCTCAGTGGAGCGCAACTGGACGATCAAGATGATGATGAAGAGACTGCCCAAA 1413
QY	1353 TCCTGAGGAATAATACTTTGATGAGCCAAATGCAGAAAGTGAATACACATATAGCAGCTC 1412
DB	1414 TCCTGAGGAATAATACTTTGATGAGCCAAATGCAGAAAGTGAATACACATATAGCAGCTC 1473
QY	1413 CTATGAAACAAATTCATTTGGTGAATTTGCCAAATGGACGACATAAAAATTTCCGAGTCAAGATT 1472
DB	1474 CTATGAAACAAATTCATTTGGTGAATTTGCCAAATGGACGACATAAAAATTTCCGAGTCAAGATT 1533
QY	1473 CCCAGAGTTTCCACCTCGTCTCTCTGATGCTTTAGAACCTGTGCGCTCGCGCTCTGT 1532
DB	1534 CCCAGAGTTTCCACCTCGTCTCTCTGATGCTTTAGAACCTGTGCGCTCGCGCTCTGT 1593
QY	1533 GCTTTCTGAGGATCACCACTTACTGAGCAAGAGGAGAGCAGTCCATCCCATGACAGAAG 1592
DB	1594 GCTTTCTGAGGATCACCACTTACTGAGCAAGAGGAGAGCAGTCCATCCCATGACAGAAG 1653

QY 1593 CAGAACGGTTTTCAGCCTCAGTACTGGGATTTGCCAAAAGCAAAAACCCCGGCGAGCTGA 1652
DB 1654 CAGAACGGTTTTCAGCCTCAGTACTGGGATTTGCCAAAAGCAAAAACCCCGGCGAGCTGA 1713
QY 1653 CTTGTTGGTGAATCCCTCGATTCCTCGGGAATGCAGATATAAATAGAGTAAAAATTTGCTGA 1712
DB 1714 CTTGTTGGTGAATCCCTCGATTCCTCGGGAATGCAGATATAAATAGAGTAAAAATTTGCTGA 1773
QY 1713 CCTGGGAATGCTTCTGCTGGTGCATATAACACTTTCACGGAAGACATCCAGAGCGCTCAGTA 1772
DB 1774 CTTGGGAATGCTTCTGCTGGTGCATATAACACTTTCACGGAAGACATCCAGAGCGCTCAGTA 1833
QY 1773 CCGCTCCATAGAGGTTTTAAATAGAGCGGGGTACAGACCCCTCGGGAATCATCTGGAGCAC 1832
DB 1834 CCGCTCCATAGAGGTTTTAAATAGAGCGGGGTACAGACCCCTCGGGAATCATCTGGAGCAC 1893
QY 1833 CGCGTGTATGGCATTTGAGCTGGCAACGGGAGATTTATTTGTTGGAACACATCTTGGGGA 1892
DB 1894 GCGGTGTATGGCATTTGAGCTGGCAACGGGAGATTTATTTGTTGGAACACATCTTGGGGA 1953
QY 1893 AGCAATATCCAGAGCGAAGACCATAGCCACATCATAGAGCTGCTAGGCGATTTCC 1952
DB 1954 AGCAATATCCAGAGCGAAGACCATAGCCACATCATAGAGCTGCTAGGCGATTTCC 2013
QY 1953 AAGGCATTTGCTCTATCTGGAATAATTTCTCGGGAATTTCTCAATCGCAGAGGAACT 2012
DB 2014 AAGGCATTTGCTCTATCTGGAATAATTTCTCGGGAATTTCTCAATCGCAGAGGAACT 2073
QY 2013 GGGACATCATCAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTTGGAAGATATGG 2072
DB 2074 GGGACATCATCAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTTGGAAGATATGG 2133
QY 2073 CTGGCCCCATGAAGTGTGCAAGTTTACAGATTTCTCGATCCCGATGTTAGAAATGGT 2132
DB 2134 CTGGCCCCATGAAGTGTGCAAGTTTACAGATTTCTCGATCCCGATGTTAGAAATGGT 2193
QY 2133 TCCAGAAAAACGAGCCTCAGCTGGCGAATGCCCTTGGCAATCTTGGTGAATTTCTTAGCA 2192
DB 2194 TCCAGAAAAACGAGCCTCAGCTGGCGAATGCCCTTGGCAATCTTGGTGAATTTCTTAGCA 2250
QY 2193 AATCTACCAATATTGCAATTTGAGCTAGCAATGTTCCAGTACATTTGGACCTAAACGG 2252
DB 2251 AATCTACCAATATTGCAATTTGAGCTAGCAATGTTCCAGTACATTTGGACCTAAACGG 2310
QY 2253 TGACTCTCAATCTTTAAACAGGATTTAAAGTGAGCTGGCTTCAATCTCAGACCTTTATTTT 2312
DB 2311 TGACTCTCAATCTTTAAACAGGATTTAAAGTGAGCTGGCTTCAATCTCAGACCTTTATTTT 2370
QY 2313 GCTTTGAGGTACTGTTGTTTGACATTTTGTGTTTGTGCACTGTGATCTCTGGGGAAGGT 2372
DB 2371 GCTTTGAGGTACTGTTGTTTGACATTTTGTGTTTGTGCACTGTGATCTCTGGGGAAGGT 2430
QY 2373 AGTCTTTGCTTCAGCTAGTATGTTTACTGACCATTTTCTTCTGGAACAATAACATGT 2432
DB 2431 AGTCTTTGCTTCAGCTAGTATGTTTACTGACCATTTTCTTCTGGAACAATAACATGT 2490
QY 2433 CTCTAAGCATGTTTCTTGTGTTGTTGACATTTCAAAATGTCATTTTGTGAATGAAAAAT 2492
DB 2491 CTCTAAGCATGTTTCTTGTGTTGTTGACATTTCAAAATGTCATTTTGTGAATGAAAAAT 2550
QY 2493 ACTTTCCCTTTGTTGTTGGCAGGTTTGTAACTATTATTAAGAAATAATTTTAGCTGA 2552
DB 2551 ACTTTCCCTTTGTTGTTGGCAGGTTTGTAACTATTATTAAGAAATAATTTTAGCTGA 2610
QY 2553 GTACTATATAATTTTCAATCTTAAGAAATTTATCAAGTTGGGAACCAAGAAATATGCAAG 2612
DB 2611 GTACTATATAATTTTCAATCTTAAGAAATTTATCAAGTTGGGAACCAAGAAATA ---CCAAG 2667
QY 2613 GAAATGTACAAATTTTATCTTCTGGCAAGGACATCATCTCTGATATATAGTGTATGTAA 2672
DB 2668 GAAATGTACAAATTTTATCTTCTGGCAAGGACATCATCTCTGATATATAGTGTATGTAA 2727
QY 2673 ATGCAACCTGTAAATGTTTACTTTGGATTAAATATGGAGGGGGGACTCAAAATTTAGAAA 2732

DB 2728 ATGCACCTGTAAATGTTACTTTCCATTAATATGGGA-GGGGGACTCAAATTTAGAAA 2786
QY 2733 AGCTAAAAA 2741
DB 2787 AGCTACCAA 2795
RESULT 9
QY15987 3708 bp DNA linear PAT 03-FEB-2004
LOCUS Sequence 1921 from Patent WO02068579.
DEFINITION QY15987
ACCESSION QY15987
VERSION QY15987.1 GI:42276844
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter,C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 1921 06-SEP-2002;
PE Corporation (US)
FEATURES
Location/Qualifiers
source 1..3708
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 76.5%; Score 2488.8; DB 6; Length 3708;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 2568; Conservative 0; Mismatches 17; Indels 9; Gaps 6;
QY 153 TCCGAAAAGCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCTCCACCGCCACC 212
DB 202 TTCGAAAAGCGGAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCTCTCCACCGCCACC 261
QY 213 ACCACCCACCGCCACCTTTTCCAGACCCACACCCCGGAGCCAGAGGAGAGATCCT 272
DB 262 ACCACCCACCGCCACCTTTTCCAGACCCACACCCCGGAGCCAGAGGAGAGATCCT 321
QY 273 GGGATCAGATGATGAGGACCAAGAGACCTTCGCGACTACTGCAAAAGGTGGATATCATCC 332
DB 322 GGGATCAGATGATGAGGAGCAAGAGACCTTCGCGACTACTGCAAAAGGTGGATATCATCC 381
QY 333 AGTGAATAATTGGAGACCTTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 392
DB 382 AGTGAATAATTGGAGACCTTTCAATGGCCGGTATCATGTTATTAGAAAGCTTGGATGGG 441
QY 393 GCACCTTCTCTACTGCTGGCTGTGCTGGGATATGAGGGGAAAAGATTTGTTGCAATGAA 452
DB 442 GCACCTTCTCTACTGCTGGCTGTGCTGGGATATGAGGGGAAAAGATTTGTTGCAATGAA 501
QY 453 AGTTGTAATAAGTGCACAGCAATTATACGAGACAGCCTTGGATGAAATAAATTTGCTCAA 512
DB 502 AGTTGTAATAAGTGCACAGCAATTATACGAGACAGCCTTGGATGAAATAAATTTGCTCAA 561
QY 513 ATGTTGTCGAGAAAGTGTATCCAGTGACCCCAACAAAGACATGTTGGTCCAGCTCATTTGA 572
DB 562 ATGTTGTCGAGAAAGTGTATCCAGTGACCCCAACAAAGACATGTTGGTCCAGCTCATTTGA 621
QY 573 CGACTTCAAGATTTTCAAGCATGAATGGGATACATGCTGTCATGGTCTTTGCAAGTACTTGG 632
DB 622 CGACTTCAAGATTTTCAAGCATGAATGGGATACATGCTGTCATGGTCTTTGCAAGTACTTGG 681
QY 633 CCAACATCTCTCAAGTGGATCATCAAAATCCAACTATCAAGCCCTCCAGTACGTTGCT 692
DB 682 CCACCATCTCTCAAGTGGATCATCAAAATCCAACTATCAAGCCCTCCAGTACGTTGCT 741

[illegible]

RESULT 10
BC068547
LOCUS
DEFINITION

BC068547 2577 bp mRNA linear PRI 06-APR-2004
Homo sapiens SFRS protein kinase 2, mRNA (cDNA clone MGC:87435
IMAGE:5265599), complete cds.

QY 573 CGACATTTCAAGATTTTCAGGATGAAATGGGATACATGTCGATGGTCTTCGAAGTACTTGG 632
DB 631 CGACATTTCAAGATTTTCAGGATGAAATGGGATACATGTCGATGGTCTTCGAAGTACTTGG 690
QY 633 CCACCATCTCTCAAGTGGATCATCAAAATCCAACTATCAAGGCTCCCACTAGCTGTTGCT 692
DB 691 CCACCATCTCTCAAGTGGATCATCAAAATCCAACTATCAAGGCTCCCACTAGCTGTTGCT 750
QY 693 GAAGAGTATCATTCGACAGGCTCTTCAAGGCTTATAGATTTACTTACACAGTAAGTGCAGAT 752
DB 751 GAAGAGTATCATTCGACAGGCTCTTCAAGGCTTATAGATTTACTTACACAGTAAGTGCAGAT 810
QY 753 CATTTCACTGACATAAAGCCGAAATATCTTGATGTTGTGATGATGATGATGATGATGATGAT 812
DB 811 CATTTCACTGACATAAAGCCGAAATATCTTGATGTTGTGATGATGATGATGATGATGATGAT 870
QY 813 AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCTCTCTCTCT 872
DB 871 AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCTCTCTCTCT 930
QY 873 TGCAGTGAATGAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCTCTCTCTCT 932
DB 931 TGCAGTGAATGAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCTCTCTCTCTCTCTCTCT 990
QY 933 AAAAAGTAAAGAAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 992
DB 991 AAAAAGTAAAGAAACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1050
QY 993 AAGAAATTTGGACGAGAGAGCTGAAAGGAAATATAGAGGAAATATAGAGGAAATATAGAGGAA 1052
DB 1051 AAGAAATTTGGACGAGAGAGCTGAAAGGAAATATAGAGGAAATATAGAGGAAATATAGAGGAA 1110
QY 1053 ACCTTCAATGACAGGATGGGAAATATCTGCCAGAGGTGAAATCTTAAAGAAACAGGATTT 1112
DB 1111 ACCTTCAATGACAGGATGGGAAATATCTGCCAGAGGTGAAATCTTAAAGAAACAGGATTT 1170
QY 1113 AGAGGAGGCGCTGAGGAGAGAGCTGCAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1172
DB 1171 AGAGGAGGCGCTGAGGAGAGAGCTGCAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1230
QY 1173 GAAAGAAAGATGCTGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1232
DB 1231 GAAAGAAAGATGCTGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290
QY 1233 TGCAGACATAGACCTTCACTGAGTGAATCACTTAAAGCAATGAGGAGGAGGAGGAGGAGGAGGAG 1292
DB 1291 TGCAGACATAGACCTTCACTGAGTGAATCACTTAAAGCAATGAGGAGGAGGAGGAGGAGGAGGAG 1350
QY 1293 CCCATTTCTCACTGGAGCAGCACTGAGCAGATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAG 1352
DB 1351 CCCATTTCTCACTGGAGCAGCACTGAGCAGATGAGTGAATGAGTGAATGAGTGAATGAGTGAATGAG 1410
QY 1353 TCTCTGAGGAAATATAATCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1412
DB 1411 TCTCTGAGGAAATATAATCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1470
QY 1413 CTATGAACAAATCAATGGTGAATGCGCAATGCGACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1472
DB 1471 CTATGAACAAATCAATGGTGAATGCGCAATGCGACATGAGGAGGAGGAGGAGGAGGAGGAGGAT 1530
QY 1473 CCCAGAGTTTCCACCTCGTTGTTCTCTGATCTTCTGATCTTCTGATCTTCTGATCTTCTGATCTTCTG 1532
DB 1531 CCCAGAGTTTCCACCTCGTTGTTCTCTGATCTTCTGATCTTCTGATCTTCTGATCTTCTGATCTTCTG 1590
QY 1533 GCTTTCTGAGGAGGATCACCCTTACTGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1592
DB 1591 GCTTTCTGAGGAGGATCACCCTTACTGAGCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1650
QY 1593 CAGAACGGTTTACGCTCCAGTACTGGGGATTTGCCAAAGCAAAACCCGGGAGCTGCA 1652
DB 1651 CAGAACGGTTTACGCTCCAGTACTGGGGATTTGCCAAAGCAAAACCCGGGAGCTGCA 1710

QY 1653 CTTGTTGGTGAATCCCTGGATCCCGGAATGCAGATAAAATTTAGAGTAAAAATTTGCTGA 1712
DB 1711 CTTGTTGGTGAATCCCTGGATCCCGGAATGCAGATAAAATTTAGAGTAAAAATTTGCTGA 1770
QY 1713 CTTGGGAAATGCTTGTGGTGCATAAACAATTCACGGAAGACATCCAGACGCGTCACTA 1772
DB 1771 CTTGGGAAATGCTTGTGGTGCATAAACAATTCACGGAAGACATCCAGACGCGTCACTA 1830
QY 1773 CCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGGAGCAC 1832
DB 1831 CCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGCACCCCTGCGGACATCTGGAGCAC 1890
QY 1833 GCGCTGTATGGCATTTGAGCTGCGCAACGCGGAGATTTATTTGTTTGAACCAATTTCTGGGGA 1892
DB 1891 GCGCTGTATGGCATTTGAGCTGCGCAACGCGGAGATTTATTTGTTTGAACCAATTTCTGGGGA 1950
QY 1893 AGACTATTCCAGAGACGAAGACCAATAGCCCAATCATATAGAGCTGCTAGGAGATTTCC 1952
DB 1951 AGACTATTCCAGAGACGAAGACCAATAGCCCAATCATATAGAGCTGCTAGGAGATTTCC 2010
QY 1953 AAGGCACCTTTGCTCTATCTGGGAAATATTTCTCGGGAATTTCTCAATTCGAGAGGAGAACT 2012
DB 2011 AAGGCACCTTTGCTCTATCTGGGAAATATTTCTCGGGAATTTCTCAATTCGAGAGGAGAACT 2070
QY 2013 GCGACACATCACAAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGAAAAATGATGG 2072
DB 2071 GCGACACATCACAAAGCTGAAGCCCTGGAGCCTCTTTGATGTACTTTGTGAAAAATGATGG 2130
QY 2073 CTGGCCCAATGAAGATGCTGCAAGTTTACAGATTTCTCGATCCCGATGTTAGAAATGTT 2132
DB 2131 CTGGCCCAATGAAGATGCTGCAAGTTTACAGATTTCTCGATCCCGATGTTAGAAATGTT 2190
QY 2133 TCAGAAAAACGAGCCTGAGCTGGGGAATGCTTGGGATCTTGGTGAATTTCTTAGCA 2192
DB 2191 TCAGAAAAACGAGCCTGAGCTGGGGAATGCTTGGGATCTTGGTGAATTTCTTAGCA 2250
QY 2193 AATTTCTACCAATATTGCAATCTGAGCTAGCAAAATGTTCCAGTACATTCGACCTAAACGG 2252
DB 2251 AATTTCTACCAATATTGCAATCTGAGCTAGCAAAATGTTCCAGTACATTCGACCTAAACGG 2310
QY 2253 TGACTCTCAATCTTTAAACAGGATTAACAGTGAAGTTCATCTCTGAGAACCTTTATTTT 2312
DB 2311 TGACTCTCAATCTTTAAACAGGATTAACAGTGAAGTTCATCTCTGAGAACCTTTATTTT 2370
QY 2313 GCTTTGAGTACTGTTGTTGATCAATTTGCTTTTGTGACATGATCTCTGGGGAAGGTT 2372
DB 2371 GCTTTGAGTACTGTTGTTGATCAATTTGCTTTTGTGACATGATCTCTGGGGAAGGTT 2430
QY 2373 AGTCTTTTGTCTTCAAGTAAAGTTCATGACCAATTTTCTTCTGGAACAATAACATGT 2432
DB 2431 AGTCTTTTGTCTTCAAGTAAAGTTCATGACCAATTTTCTTCTGGAACAATAACATGT 2490
QY 2433 CTCTAAGCATTTGTTCTTGTGTTGTGATCAATTCAAATGTCATTTTTTTTGAATGAAAAAT 2492
DB 2491 CTCTAAGCATTTGTTCTTGTGTTGTGATCAATTCAAATGTCATTTTTTTTGAATGAAAAAT 2550
QY 2493 ACTTTCCCTCT 2502
DB 2551 ACTTTCCCTCT 2560

RESULT 11

BC020178
LOCUS
DEFINITION
Mus musculus serine/arginine-rich protein specific kinase 2, mRNA
(cDNA clone MGC:27638 IMAGE:4507346), complete cds.
ACCESSION
BC020178
VERSION
BC020178.1
KEYWORDS
MGC.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length

human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 4287)

Direct Submission

Strausberg,R.

Submitted (26-NOV-2003) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs@mail.nih.gov

Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library Preparation: M. Bento Soares, University of Iowa

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: University of Iowa, Dr. M. Bento Soares and Dr.

Thomas L. Casavant.

Web site: <http://genome.uiowa.edu>

Contact: bento-soares@uiowa.edu; tom-casavant@uiowa.edu

Bonaldo,M.F., Akabogu,I., Bair,T., Bair,J., Crouch,K., Davis,A.,

Fishler,K., Keppel,C., Kucaba,T., Lebeck,M., Melo,A., Schaefer,K.,

Scheetz,T., Smith,C., Snir,E., Tack,D., Trout,K., Walters,J.,

Casavant,T., Soares,M.B.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: Plate: Row: Column: 0

This clone has the following problem: The cds is short compared to

the longest cds in the locus.

Location/Qualifiers

1. .4287

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:684273"

/tissue_type="Brain"

/clone_lib="NIH_BMAP_CH0"

/lab_host="DH10B"

/note="Vector: pYX-ASC"

1. .4287

/gene="Srpk2"

/note="synonyms: WRP6, msRPK2"

/db_xref="LocusID:20817"

/db_xref="MGI:1201408"

447. .917

/gene="Srpk3"

/codon_start=1

/product="Srpk2 protein"

/protein_id="AAH62941.1"

/db_xref="GI:38566030"

/db_xref="LocusID:20817"

/db_xref="MGI:1201408"

/translation="WGKRFVAMKVKSQHYETALDEIKLLKCVRESDDPNKDM

VVQLDDFKISGMNGHVMVEVLGHLLKWLKSNYQGLPVRCKSIIRVQLGLD

YHSKCKIHTDKPENILMVCDDAYVRMAEATWQXAGAPPPSGSADRKN1"

456. .812

/gene="Srpk2"

/note="S TK; Region: Serine/Threonine protein kinases,

catalytic domain. Phosphotransferases of the serine or

threonine-specific kinase subfamily. The enzymatic

activity of these protein kinases is controlled by

phosphorylation of specific residues in the activation

misc_feature

segment of the catalytic domain, sometimes combined with
reversible conformational changes in the C-terminal
autoregulatory tail"
/db_xref="CDD:cd00180"

ORIGIN

Query Match 61.5%; Score 2001.4; DB 10; Length 4287;
Best Local Similarity 89.0%; Pred. No. 0;
Matches 2286; Conservative 0; Mismatches 221; Indels 62; Gaps 9;

QY 106 GTGCTGGCCATTTCAGGCGCCGGAAGCGGAGCGGCGGAAAGAGAGAAACATCCCGAAAGGCGG 165

Db 135 GTGCTGGCCATTTCAGGCGCCGGAAGCGGAGCGGCGGAAAGAGAGAAACATCCCGAAAGGCGG 194

QY 166 GAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCCACCGCCACACACACACACACCG 225

Db 195 GAGCCTCAACAGAAAGCTCTTTAGTTCTCTCTCCACCGCCACACACACACACCG 254

QY 226 CCACCTTTGCGAGACCCCGACACCCCGGAGCCAGAGGAGAGATCCTGGGATCAGATGAT 285

Db 255 C-----TGGCAGACCCCGACACCCCGGAGCCAGAGGAGAGATCCTGGGATCAGATGAT 308

QY 286 GAGGAGCAAGAGGACCTCGCGACTACTGCAAAAGTGGATATCATCCAGTCAAAATTGGA 345

Db 309 GAGGAGCAAGAGGACCTCGCGACTACTGCAAAAGTGGATATCATCCAGTCAAAATTGGA 368

QY 346 GACCTCTTCAATGGCGGGTATCATGTTATTAGAAAGCTTGGATGGGCGACCTTCTTACT 405

Db 369 GATCTCTTCAATGGCGGGTATCATGTTATTAGAAAGCTTGGATGGGCGACCTTCTTACT 428

QY 406 GTCTGGCTGTCTGGGATATCGAGGGAAGAGATTTGTTGCAATGAAAGTTGTAAGAGT 465

Db 429 GTATGGCTGTCTGGGATATCGAGGGAAGAGATTTGTTGCAATGAAAGTTGTAAGAGT 488

QY 466 GCCAGCATTTATCGGAGACAGCCTTGATGCAATGCAATGCTCAAAATGTTTTCAGAA 525

Db 489 GCCAGCATTTATCGGAGACAGCCTTGATGCAATGCAATGCTCAAAATGTTTTCAGAA 548

QY 526 AGTGATCCAGTGACCCAAACAGACATGGTGTGTCAGCTCATTTGACGATTTCAAGATT 585

Db 549 AGTGATCCAGTGACCCAAACAGACATGGTGTGTCAGCTCATTTGACGATTTCAAGATT 608

QY 586 TCAGGCATGATGGGATACATGTCGATGTCCTCGAAGTACTTGGCCACCATCTCTC 645

Db 609 TCAGGCATGATGGGATACATGTCGATGTCCTCGAAGTACTTGGCCACCATCTCTC 668

QY 646 AAGTGGATCATCAATCCAACTATCAAGGCTCCAGTACGTTGTGTGAAGAGTATCATTT 705

Db 669 AAGTGGATCATCAATCCAACTATCAAGGCTCCAGTACGTTGTGTGAAGAGTATCATTT 728

QY 706 CGACAGGTCTTTCAAGGGTTAGATTATTTACAGTAAGTCAAGATCATTTCAATCTGAC 765

Db 729 CGACAGGTCTTTCAAGGGTTAGATTATTTACAGTAAGTCAAGATCATTTCAATCTGAC 788

QY 766 ATAAAGCGGAAATATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 825

Db 789 ATAAAGCGGAAATATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 848

QY 826 GAGGCACTGAGTGGCAGAAAGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 885

Db 849 GAGGCACTGAGTGGCAGAAAGCAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900

QY 886 GCTCCACGACAGAAACCTTATAGGAAAAATATCTAAAAAACAAGAAAAAATCTGAAAG 945

Db 901 -----ATAGAAAAATATCTAAAAAACAAGAAAAAATCTGAAAG 942

QY 946 AAACAGAGAGCAGGCTGAGTTATTGGAGAGCCCTGCGAGAGATGAGAGAAATTTGGAG 1005

Db 943 AAACAGAGAGCAGGCTGAGTTATTGGAGAGCCCTGCGAGAGATGAGAGAAATTTGGAG 1002

QY 1006 CGAGAGCTGAAAGGAAAAATATAGAGAAAAACATCAGCTCAGCTGCGACCTTCC-----AAT 1062

Db 1003 CGAGAGCTGAAAGGAAAAATATAGAGAAAAACATCAGCTCAGCTGCGAGAGCTTCCGGGGAG 1062

GKRFVAMKVISAQHYTTALDEIKLLKCVRESDPSPDNKDMVQLIDDFKISGMNGI
 NLCMVFEVLGHLLKWIITKSNYQGLPVRVCSIIIRVQLGLDYLHSHKSIITHDIKPE
 HILMCDLHYVRRWAAEATEWQKAGAPPSGAVSTAPOQPIGKISNKKKGLKKQK
 KROALLKRLQETEEELREARERKIEENITSAAPSDODGEYCEPVKLKTTGLEEFA
 EAETAKNGEAEDEKEDEKENIEKDEDVDDELANIIDPTWIESPKTNHGIENPFP
 SLEQQLDDEDDDEDCNPEYNLDEPNAESDYTYSSVEQFNGELPNRHKIPESOF
 PEFSTSLFSGSLPEPVACSGVLSBESPLTEQSEBSPSHDRSRTVSASSTGDLPLKAKTRA
 ADLLNPLDPRNADKIRVKIADLGNACWVQKPGQLTCW"

ORIGIN

Query Match 59.9%; Score 1949.8; DB 9; Length 3028;

Best Local Similarity 91.5%; Pred. No. 0;

Matches 2168; Conservative 0; Mismatches 7; Indels 194; Gaps 2;

Qy	153	TCGGAAGAGCGGAGCCTCAACAGAAAGCTCTTTAGTTCCTCTCTCCACCGCCACC	212
Db	252	TTGCAAGAGCGGAGCCTCAACAGAAAGCTCTTTAGTTCCTCTCTCCACCGCCACC	311
Qy	213	ACCACACACCGCCACCTTTGCCAGACCCCHACCCCGGAGCCAGAGGAGATCCT	272
Db	312	ACCACACACCGCCACCTTTGCCAGACCCCHACCCCGGAGCCAGAGGAGATCCT	371
Qy	273	GGGATCAGATGATGAGGAGCAAGAGACCTCGCGACTACTGCAAGAGGTGGATATCATCC	332
Db	372	GGGATCAGATGATGAGGAGCAAGAGACCTCGCGACTACTGCAAGAGGTGGATATCATCC	431
Qy	333	AGTGAAATTTGAGACCTCTTCAATGGCGGTATCATGTATTTAGAAAGCTTGGATGGG	392
Db	432	AGTGAAATTTGAGACCTCTTCAATGGCGGTATCATGTATTTAGAAAGCTTGGATGGG	491
Qy	393	GCATTTCTACTGTCTGCTGTGGATATGAGGGGAAAGATTTGTTGCAATGAA	452
Db	492	GCATTTCTACTGTCTGCTGTGGATATGAGGGGAAAGATTTGTTGCAATGAA	551
Qy	453	AGTTGTAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAAATAAAATTCCTCAA	512
Db	552	AGTTGTAAAGTGGCCAGCATTTATACGGAGACAGCCTTTGGATGAAATAAAATTCCTCAA	611
Qy	513	ATGTGTTTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGTGGTCCAGCTCATTTGA	572
Db	612	ATGTGTTTCGAGAAAGTGATCCAGTGACCCAAACAAAGACATGTGGTCCAGCTCATTTGA	671
Qy	573	CGACTTCAAGATTTGAGGATGAATGGGATACATGTCTGCATGGTCTTCGAAGTACTTGG	632
Db	672	CGACTTCAAGATTTGAGGATGAATGGGATACATGTCTGCATGGTCTTCGAAGTACTTGG	731
Qy	633	CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCTCCAGTACGTTGTGT	692
Db	732	CCACCATCTCCTCAAGTGGATCATCAAAATCCAACTATCAAGGCCTCCAGTACGTTGTGT	791
Qy	693	GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTACTTACACAGTAAGTCAAGAT	752
Db	792	GAAGAGTATCATTCGACAGGTCTTCAAGGGTTAGATTACTTACACAGTAAGTCAAGAT	851
Qy	753	CATTCTACTGACATTAAGCCGGAATATCTTGATGTGTGGATGATGATATGTGAG	812
Db	852	CATTCTACTGACATTAAGCCGGAATATCTTGATGTGTGGATGATGATATGTGAG	911
Qy	813	AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCCTTCCAGGGTC	872
Db	912	AAGAAATGGCAGCTGAGGCCACTGAGTGGCAGAAAGCAGGTGCTCCTCCTTCCAGGGTC	971
Qy	873	TGCAGTGTAGTACGGTCCACAGCAGAAACCTTAGGAAATAATCTTAAACAAAGAA	932
Db	972	TGCAGTGTAGTACGGTCCACAGCAGAAACCTTAGGAAATAATCTTAAACAAAGAA	1031
Qy	933	AAACTGAAAGAAACAGAGGAGCGGTGATTTTGGAGAGCGCCTCGAGAGAT	992
Db	1032	AAACTGAAAGAAACAGAGGAGCGGTGATTTTGGAGAGCGCCTCGAGAGAT	1091
Qy	993	AGAAGAAATTTGGAGCGAGAGCTGAAAGGAAATAATAGAGAAACATCACCTCGAGTGC	1052
Db	1092	AGAAGAAATTTGGAGCGAGAGCTGAAAGGAAATAATAGAGAAACATCACCTCGAGTGC	1151

Qy	1053	ACCTTCCAAATGACACAGGATGGCGAATACTGCCAGAGAGTGAACTAAACAAACACAGGAT	1112
Db	1152	ACCTTCCAAATGACACAGGATGGCGAATACTGCCAGAGAGTGAACTAAACAAACACAGGAT	1211
Qy	1113	AGAGGAGCGGCTCAGGACAGAGCTGCAAGAGCAATGGTGAAGCTGAGGACACAGGAAGA	1172
Db	1212	AGAGGAGCGGCTCAGGACAGAGCTGCAAGAGCAATGGTGAAGCTGAGGACACAGGAAGA	1271
Qy	1173	GAAAGAAAGATGCTGAGAAAGAAAAATTGAAAAAGATGAAGATGATGTAGATCAGGAAT	1232
Db	1272	GAAAGAAAGATGCTGAGAAAGAAAAATTGAAAAAGATGAAGATGATGTAGATCAGGAAT	1331
Qy	1233	TGCGAAACATGAGACCTTACGTGGATAGAAATCACTTAAACCAATGCCATATTTGAGATGG	1292
Db	1332	TGCGAAACATGAGACCTTACGTGGATAGAAATCACTTAAACCAATGCCATATTTGAGATGG	1391
Qy	1293	CCCATTTCTACTGAGCAGCAACTGGACGATGAAGATGATGATGAAGAGACTGCCCAAA	1352
Db	1392	CCCATTTCTACTGAGCAGCAACTGGACGATGAAGATGATGATGAAGAGACTGCCCAAA	1451
Qy	1353	TCCTGAGGAATATTAATCTTGATGAGCCAAATGCAAGAGTGAATTACACATATAGCAGTC	1412
Db	1452	TCCTGAGGAATATTAATCTTGATGAGCCAAATGCAAGAGTGAATTACACATATAGCAGTC	1511
Qy	1413	CTATGAACAATTTCAATGGTGAATTTGCCAAATGGACGACATAAATTTCCGAGTACAGTT	1472
Db	1512	CTATGAACAATTTCAATGGTGAATTTGCCAAATGGACGACATAAATTTCCGAGTACAGTT	1571
Qy	1473	CCGAGAGTTTTCCACCTCTGTTGTTCTCTGGATCCTTAGAACCTGTGGCTCGGCTCTGT	1532
Db	1572	CCGAGAGTTTTCCACCTCTGTTGTTCTCTGGATCCTTAGAACCTGTGGCTCGGCTCTGT	1631
Qy	1533	GCTTTCTGAGGGATCACCACTTACTGACAAAGAGAGAGCAGTCCATCCCATGACAGAAG	1592
Db	1632	GCTTTCTGAGGGATCACCACTTACTGACAAAGAGAGAGCAGTCCATCCCATGACAGAAG	1691
Qy	1593	CAGAACGGTTTTGAGCTCCAGTACTGGGATTTGCCAAA	1632
Db	1692	CAGAACGGTTTTGAGCTCCAGTACTGGGATTTGCCAAAAGCAAAACCCGGGACAGTGA	1751
Qy	1633	-----	1632
Db	1752	CTTGTGTGTAATCCCCTGGATCCGGAATGCAGATAAAATTAGAGTAAATAATTGCTGA	1811
Qy	1633	-----GCAAAACCCGGGAGCTGACTTGTGGTGAATCCCTCTG	1671
Db	1812	CCTGGGAAATGCTTGTGGTGCAAAACCCGGGAGCTGACTTGTGGTGAATCCCTCTG	1871
Qy	1672	GATCCGCGGAATGCAGATAAAATTAGAGTAAATAATTGCTGACCTGGGAAATGCTTGTGG	1731
Db	1872	GATCCGCGGAATGCAGATAAAATTAGAGTAAATAATTGCTGACCTGGGAAATGCTTGTGG	1931
Qy	1732	GTGCATAAACACTTCCACGGAAGACATCCAGACCGGTCTAGTACCGCTCCATAGAGTTTTA	1791
Db	1932	GTGCATAAACACTTCCACGGAAGACATCCAGACCGGTCTAGTACCGCTCCATAGAGTTTTA	1991
Qy	1792	ATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAGACCGCGTGTATGGCATTGAG	1851
Db	1992	ATAGAGCGGGGTACAGCACCCCTCGGACATCTGGAGACCGCGTGTATGGCATTGAG	2051
Qy	1852	CTGGCAAACGGGAGATTTTGTGTAACCACTCTGGGAGAGACTATTTCCAGAGACGAA	1911
Db	2052	CTGGCAAACGGGAGATTTTGTGTAACCACTCTGGGAGAGACTATTTCCAGAGACGAA	2111
Qy	1912	GACCACATAGCCCATCATAGAGCTGTAGGAGTATTTCCAAAGGCACTTTGCTCTATCT	1971
Db	2112	GACCACATAGCCCATCATAGAGCTGTAGGAGTATTTCCAAAGGCACTTTGCTCTATCT	2171
Qy	1972	GGAAATAATTTCTGGGAATTTCTCAATCGCAG	2003
Db	2172	GGAAATAATTTCTGGGAATTTCTCAATCGCAGAGATCATATAGCATTGATCATTTGAACCTG	2231

```
QY 2004 ----- 2003
Db 2232 CTGGGGAAGTCCCTCGAATAACGCTATGTTGGGGAATATCTCCAAGGAGTTCCTCAC 2291
QY 2004 -----AGGAGAACTCGGACACATACCAAGCTGAAGCCCTGGAGCCTCTTTGATGACTT 2058
Db 2292 AGAAAAGGAGAACTCGGACACATACCAAGCTGAAGCCCTGGAGCCTCTTTGATGACTT 2351
QY 2059 GTGGAAGATGCTGGGCCCACTGAAGATGCTGCAAGTTTACAGATTTCTCTGATCCCG 2118
Db 2352 GTGGAAGATGCTGGGCCCACTGAAGATGCTGCAAGTTTACAGATTTCTCTGATCCCG 2411
QY 2119 ATGTTAGAAATGTTCCAGAAAACGAGCTCAGCTGGGGAATGCCCTCGGCATCCTTGG 2178
Db 2412 ATGTTAGAAATGTTCCAGAAAACGAGCTCAGCTGGGGAATGCCCTCGGCATCCTTGG 2471
QY 2179 TTGAATTTCTAGCAATTTCTACCAATATTGCAATTTCTGAGCTAGCAAAATGTTCCAGTACA 2238
Db 2472 TTGAATTTCTAGCAATTTCTACCAATATTGCAATTTCTGAGCTAGTAAATGTTCCAGTACA 2531
QY 2239 TTGGACCTAAACGCTGACTCTCATTTCTTTAAACGAGTTTACAAGTGAGCTGGCTTCATCT 2298
Db 2532 TTGGACCTAAACGCTGACTCTCATTTCTTTAAACGAGTTTACAAGTGAGCTGGCTTCATCT 2591
QY 2299 CAGACCTTTATTTGCTTTGAGTACTGT 2327
Db 2592 CAGACCTTTATTTGCTTTGAGTTTTGT 2620

RESULT 14
AB006036
LOCUS AB006036 2300 bp mRNA linear ROD 25-DEC-1997
DEFINITION Mus musculus mRNA for SRPK2, complete cds.
ACCESSION AB006036
VERSION AB006036.1 GI:2723281
KEYWORDS SRPK2.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Kuroyanagi.N., Onogi.H., Wakabayashi.T. and Hagiwara.M.
TITLE Novel SR-protein-specific kinase, SRPK2, disassembles nuclear
speckles
Biochem. Biophys. Res. Commun. (1998) In press
2 (bases 1 to 2300)
Hagiwara.M.
Direct Submission
Submitted (25-JUL-1997) Masatoshi Hagiwara, Tokyo Medical and
Dental University Medical Research Institute, Department of
Endocrinology; Yushima, Bunkyo-ku, Tokyo 113, Japan
(E-mail:m.hagiwara.end@mri.tmd.ac.jp, Tel:03-5803-5836,
Fax:03-5803-5836)
Sequence updated (16-DEC-1997).
Location/Qualifiers
1..2300
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..2046
/codon_start=1
/product="SRPK2"
/protein_id="BAA24055.1"
/db_xref="GI:2723282"
/translation="MSVNSEKSSSRPEPQOKAPLVPPPPPPPPPLPDPAPEPE
BEILGSDDEQDPADYCKGYHPVKIGLDFNGRHHVIRKLGHSFTWLCHWDQGK
RFVAMKVSAOHYETETALDEIKLKRVRESDFPNKDMVQLDDDFKISGNGIHW
CMVFELVHLGDKWIIKSNYQGLPVCVKSIIRQLGLDLYHSCKIIHTDKPENI
LMCDDVAYRRMAATEATQWAGAPPSSAVSTAPQPIGKI SKNKKKKLKKQKR
QAELEKRLQIETELREAREARKILRENTSASGEODGEYQPEVTLKAADLEDTEE
ETAKNGEVEDDEKEDAEKNAEKDEDDVEOPANLDPPTWVESPKANGHIENGPSL
EQKNGEEDDDCANPEYNDLDEPNAESDITYSSSYEQPNGLPNQHKITSEPTPL
FSGPLEPVACGSVI SEGSLTEQESSSPSHDRSRTVSASSTGDLPTKTKRAADLLVNP
```

```
LDPRNADKIRVKIADLGNACWVHKHFTEDIQTRQYRSIEVLIGAGYSTPADIWSTACM
APELATGYDLFPESHGEDIYRDIEDIAHIEELGSIIPRPFALSGKTSRFFNRRGEIR
HTTKLKPSWLFDFVLVEKVGWPHEDAAQFTDFLIPMLEMVPKRSAGECLRHFWLNS"
2300
/note="18 A nucleotides"

polyA_site
ORIGIN
Query Match 57.2%; Score 1861.4; DB 10; Length 2300;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 2054; Conservative 0; Mismatches 216; Indels 23; Gaps 4;

QY 153 TCCGAAAAGAGCCGAGCCTCAACAGAAAAGCTCTCTTTAGTTCTCTCTCTCCACCGCAC 212
Db 30 TTCAGAAAGCCGAGCCTCAACAGAAAGCTCTCTTTAGTTCTCTCTCTCCACCGCAC 89
QY 213 ACCACACACACCCACCTTTGCCAGACCCACACCCCGAGCCAGAGGAGAGATCCT 272
Db 90 ACCGCCACCCACAC-----TGCCAGACCCCGCGCCCCAGAGCCAGAGGAGATCT 143
QY 273 GGGATCAGATGATGAGGACCAAGAGACCCCTCGGACTACTGCAAAAGTGGATATCATCC 332
Db 144 GGGGTGAGATGATGAGGACCAAGAGACCCCGCAGATTACTGCAAAAGTGGATATCATCC 203
QY 333 AGTGAATAATTGAGACCTCTTCAATGGCCGGTATCATGTTATTAGAAAAGCTTTGGTGA 392
Db 204 AGTGAATAATTGAGATCTCTTCAATGGTGCATATCATGTCATTAGAAAGTGGATGGG 263
QY 393 GCACCTTCTACTGTCTGCTGCTGGGATATGCAGGGGAAAAGATTTTGGCAATGA 452
Db 264 GCACCTTCTACTGTATGCTGCTGGGATATGCAAGGAAAAGATTTTGGCAATGA 323
QY 453 AGTTGTAAAAGTGGCCAGCATTTATCGGAGACAGCCTTGGATGAAATAAATTTGCTCAA 512
Db 324 AGTTGTAAAAGTGGCCAGCATTTATACAGAGACAGCCTTGGATGAAATAAATTTGCTCAA 383
QY 513 ATGCTGTTGAGAAAGTATCCAGTGACCCCAACAAAGACATGGTGGTCCAGCTCATGA 572
Db 384 ATGCTGTTGAGAAAGTATCCAGTGACCCCAACAAAGACATGGTGGTGGTAAATTTGA 443
QY 573 CGACTTCAAGATTTTCAGGCATGAATGGGATACATGCTGTCATGCTTCCAGAGTACTTG 632
Db 444 TGACTTCAAGATCTCAGGCATGAATGGGATACATGCTGTCATGCTTTCAGAGTACTTG 503
QY 633 CCACATCTCTCAAGTGGATCATCAAAATCAAACTATCAAGGCCTCCAGTACTGTTGT 692
Db 504 TCACCATCTCTCAAAATGGATCATCAAAATCAAACTATCAAGGCCTCCAGTACTGTTGT 563
QY 693 GAAGAGTATCATTCGACAGGTCCTTCAGGGTTAGATTACTTACACAGTAAGTGCAAGAT 752
Db 564 GAAGAGTATCATTCGACAGGTCCTTCAGGGTTAGATTATCTACACAGTAAGTGCAAGAT 623
QY 753 CATTTCATCTGACATAAAGCCCGAAAATATCTTTGATGCTGTTGGATGATGCATATGTG 812
Db 624 AATTTCACACCGACATAAAGCCCGAAAATATCTTTGATGCTGTTGGATGATGCATGTG 683
QY 813 AAGAAATGGCAGTGGAGCCACTTGAGTGGCAGAAAAGCAGGTGCTCTCTCTTCAGGGTC 872
Db 684 AAGAAATGGCAGCAGGAGCCAGCAGAGTGGCAGAAAAGCAGGTGCTCTCTCTTCAGGGTC 743
QY 873 TGCAGTGGTACGGCTCCACAGCAGAAACCTTAGGAAAATATCTTAAAAACAAAAGAA 932
Db 744 TGCAGTGGATGACGGCTCCACAGCAAAACCTTAGGAAAATATCTTAAAAACAAAAGAA 803
QY 933 AAAAAGTGAAGAAAAGAACAGAGAGCGGCTGAGTTATTGGAGAACGGCTGTCAGGAGAT 992
Db 804 AAAGCTGAAAAGAAAAGAACAGAGAGCAGGCTGAGTTGCTGGAGAAACGGCTACAGGAT 863
QY 993 AGAAGAATTTGGAGCAGAGAGCTGAAAGGAAAAATAATAGAAAGAAAACATCACCTCAGTC 1052
Db 864 TGAGGAAATTTGGAGCAGAGAGCGAAGGAAAAATCTTAGAGGAGAAACATCACCTCTCAGA 923
QY 1053 ACCTTCCAAATGACAGGATGGCGAATACTGTCAGGAGTGAACCTTAAAAACAAACAGGAT 1112
Db 1112
```

Db 924 AGCTTCCGGGAGCAGGATGGAGAGTACGAGCCGAGGTGACATGAAAGCAGCCGACTT 983
QY 1113 AGAGAGGCGGCTGAGGCGAGAGACTGCAAGGACAAATGGTGAAGCTGAGGACACAGGAAGA 1172
Db 984 AGAGGACACAACTGAGGAGAGAGACAGCAAGAGATATGGTGAAGTTGAAGACAGGAAGA 1043
QY 1173 GAAAGAAGATGCTGAGAAAGAAAACATTGAAAGAGATGAAGATGATGTGAATCAGGAAGT 1232
Db 1044 GAAAGAAGATGCGGAGAGAGAACCGCGAGAGAGATGAAGATGATGTGAACAGGAAGC 1103
QY 1233 TCGCAACATAGACCCCTACGCTGATAGATCACTAAACCAATGCGCATATTTGAGAAATGG 1292
Db 1104 TGCAAACTTAGACCCGACCTGGGTGGAGTCCCGAAAGCCAATGGCCATATTTGAAATGG 1163
QY 1293 CCCATTCTCAGTGGAGCAGCAACTGGACGATGAAGATGATGAAGAGACTGCGCCAAA 1352
Db 1164 CCGCTTCTCAGTGGAGCAGCAGCTGGAGGATGAAGAGCAGATGAAGATGACTGTGCAAA 1223
QY 1353 TCTTGAGGAATATAATCTTTGATGAGCCAAATGCAAGAAAGTATTACATATAGCAGCTC 1412
Db 1224 TCCCGAGGAGTATAACCTCGATGAGCCAAATGCAGAGAGTATTACAGTATAGCAGCTC 1283
QY 1413 CTATGAACAAATCAATGTGTGAATTCGCAAAATGGCAACATTAAGACT----- 1329
Db 1284 CTATGAACAAATCAATGTGTGAATTCGCAAAATGGCAACATTAAGACT----- 1329
QY 1473 CCAGAGTTTCCACCTCGTCTGCTCTCTGATCCTTAGAACCTGTGGCTGGGCTCTGT 1532
Db 1330 -TCAGATTTCCACACCGTTGTTCTGGGCGCTTAGAACCTGTGGCTGGGCTCTGT 1388
QY 1533 GCTTTCTGAGGATCACTTACTGAGCAGAGAGAGCAGTCCATCCATGACAGAG 1592
Db 1389 GATTTCAAGAGGATGCCACTTACCAGCAGGAGGAGAAAGCAGTCCCTCCATGACAGAG 1448
QY 1593 CAGACGGTTTCAGCTCAGTACTGGGATTTGCCAAGCAAAAACCGGCGAGCTGA 1652
Db 1449 CAGGACAGTTTCAGCTCAGTACTGGGATTTGCCAAGCAAAAACCGGCGGCTGA 1508
QY 1653 CTGTTGGTGAATCCCTGGATCCGCGAATCAGATAAAATTAGAGTAAAAATTCCTGA 1712
Db 1509 CTGTTGGTGAACCTCTGGATCCAGGATGAGATAAAATTAGAGTAAAAATTCCTGA 1568
QY 1713 CTGGAATATGTTGTTGGTGCATAAACAATTCACGGAAGACATCCAGACCGCTGAGTA 1772
Db 1569 CTTGGAAATGCTTTGTTGGTGCATAAACAATTCACAGAGGATATCCAGACACGTCAGTA 1628
QY 1773 CCGCTCCATAGAGTTTAAATAGGAGCGGGTACAGCACCCCTGGGACATCTGGAGCAC 1832
Db 1629 TAGGTCCATAGAGTTTGTATAGGAGCAGGCTACAGCACCTCGAGATATTTGGAGTAC 1688
QY 1833 GCGGTGATGGCATTGAGCTGGCAACGGGAGATTAATTTGTTGAACACATTCCTGGGA 1892
Db 1689 AGCTTGATGGCATTGAGCTGGCCACAGAGACTATTTGTTGCAACCCGATCTGGGA 1748
QY 1893 AGACTATTCAGAGACGAAGACACATAGCCACATCATAGAGCTGTAGGACGATATCC 1952
Db 1749 AGACTATTCAGAGATGAAGACACATAGCCACATCATAGAGCTGTAGGACGATATCC 1808
QY 1953 AAGGCACTTGTCTATCTGGAATAATTTCTGGGAATTTCTCAATCGCAGAGGAGACT 2012
Db 1809 AAGGCACTTGTCTGTCTGGAATAATTTCTGGGAATTTCTCAATCGCAGAGGAGACT 1868
QY 2013 GCGACACATCACCAGCTGAAGCCCTGGAGCCTCTTTGATGACTTTGTGGAAGATGAG 2072
Db 1869 GCGGACATCACCAGCTGAAGCCCTGGAGCCTCTTTGATGACTTTGTGGAAGATGAG 1928
QY 2073 CTGGCCCCATGAGATGCTGCAAGTTTACAGATTTCTCGATCCCGATTTGAAATGTT 2132
Db 1929 CTGGCCCCATGAGATGCTGCAAAATTTACAGATTTCTCGATCCCGATTTGAGATGTT 1988
QY 2133 TCCAGAAAACAGCCTCAGCTGGCGAAGTCCCTGGGATCTTTGGTGAATTTCTTAGCA 2192
Db 1989 TCCAGAAAACAGCCTCAGCTGGCGAATTCCTTTGCACATCTCTGTTGAATTTCTTAGCA 2048

QY 2193 AATTCTACCAATATTGCATTTCTGAGCTAGCAAAATGT-TCCAGTACATTTGACCTAAACG 2251
Db 2049 GATTCTCAAAATATAGCATTTCTGAGCTAGCAAAATTTCTCCAGTACATTTGACCTACACA 2108
QY 2252 GTGACTCTCATCTTTTAACAGGATTACAGTGAAGTGGCTTCATCTCAGACCTTTATTT 2311
Db 2109 GTGGCTCTCATTTCTTTAACAGGATTACAGTGAAGTGGCTTCATCTCAGACCTTTTGTAG 2168
QY 2312 TGCCTTGAGGTACTGTTGTTTGTGACATTTGCTTTTGTGCACTGTGATCCCTGGGAAGGG 2371
Db 2169 TGCCTTGAGGTACTGTTGTTTGTGACATTTGCTTTCTGTGCACTGTGCTTGGGAGGC 2228
QY 2372 TAGTCTTTTGTCTTCACTAAGTAGTATTACTGACCATTTCTTTGGAACAATAAATG 2431
Db 2229 TGGTCT-TTTGCTTTCAGCTAAGTAGTTTACTGGCCATTTCTTCTGGAACAATAAATG 2287
QY 2432 TCTCTAAGCATTTG 2444
Db 2288 TCTCTAAGCATTTG 2300

RESULT 15

AK093420

LOCUS

DEFINITION

Homo sapiens cdna FLJ36101 fis, clone TEST12021118, highly similar

to Mus musculus mRNA for SRPK2.

ACCESSION

AK093420.1 GI:21752284

VERSION

oligo capping; fis (full insert sequence).

KEYWORDS

Homo sapiens (human)

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,
Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,
Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A.,
Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imobe, N., Musashino, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,
Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Sato, N., Takami, S.,
Terasahima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikemura, Y., Okamoto, S., Okitani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, S., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.

Complete sequencing and characterization of 21,243 full-length

human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

JOURNAL

PUBMED

REFERENCE

AUTHORS

Nishi, T., Ota, T., Nakagawa, S., Senoh, A., Mizuguchi, H., Inagaki, H.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,

Mon Jan 24 10:28:30 2005

QY	2584	TCAAGTTGGGACCAAGAAATAGCAAGGAAATGTACAATTTTATCTTCTGGCAAGGG	2643
Db	2417	TCAAGTTGGGACCAAGAAATA---GCAAGGAAATGTACAATTTTATCTTCTGGTAAAGGG	2473
QY	2644	ACATCATTCCTGTATTATAGTGTATGTAATGCACCCCTGTAAATGTTACTTTGGATTAAA	2703
Db	2474	ACATCATTCCTGTATTATAGTGTATGTAATGCACCCCTGTAAATGTTACTTTCCATTAAA	2533
QY	2704	TATGGAGGGGGGACTCAAAATTCAGAAAAGCT	2736
Db	2534	TATGGGA-GGGGGACTCAAAATTCAGAAAAGCT	2565

Search completed: January 23, 2005, 17:20:50
Job time : 13476 secs